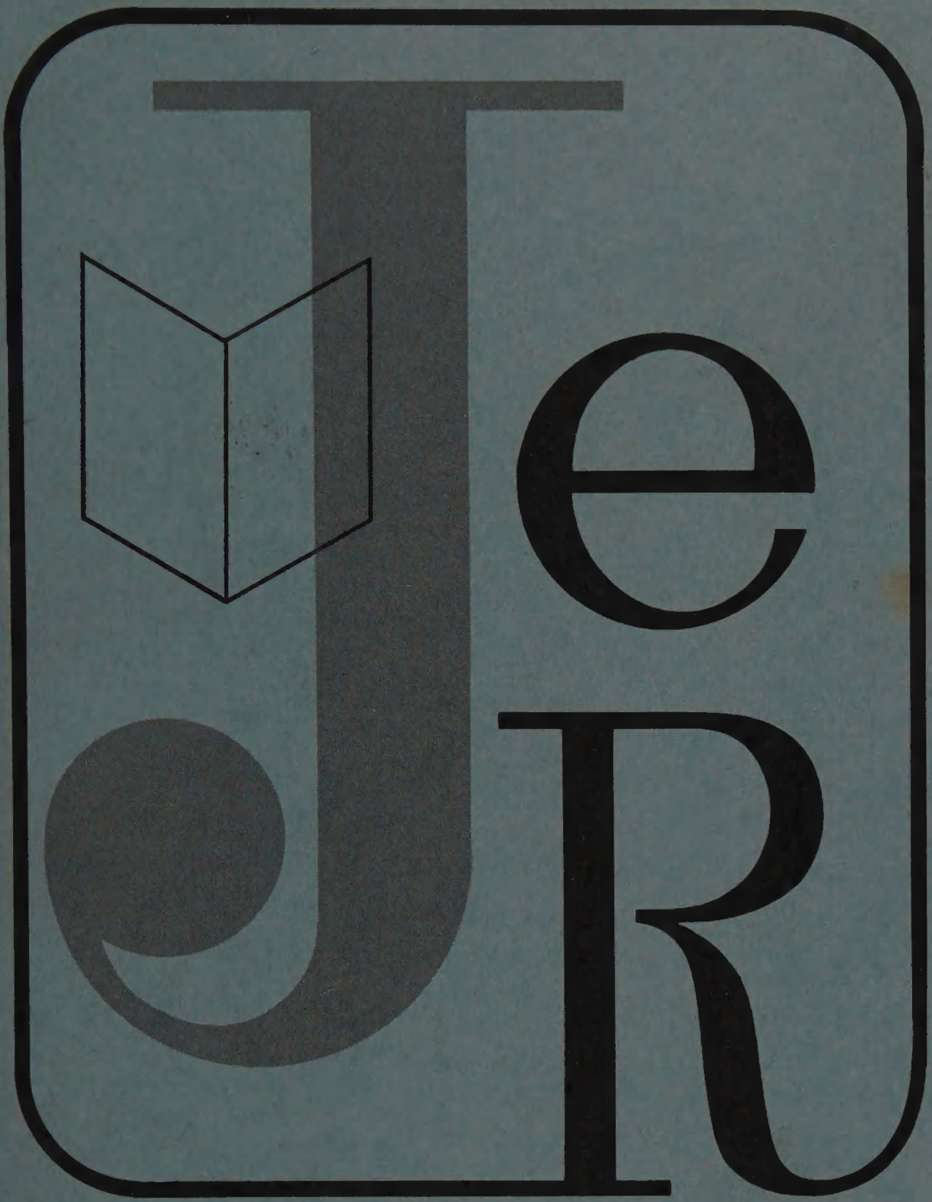
Rank Order, Shepps

The Journal of Educational Research

Dedicated to the Study of Education



Volume 65 OCTOBER 1971 Number 2

Policy and Perspective

by **LINDLEY J. STILES**

Professor of Education for Interdisciplinary Studies, Sociology, and Political Science, Northwestern University

Public Support For Nonpublic Schools

STRONG POLITICAL campaigns are generating in state after state to enact appropriations for nonpublic schools. The proposals range from funds for textbooks, transportation, scholarships, tuition vouchers, purchase of services, and teachers' salaries to unrestricted per-pupil grants. Support comes from the powerful parochial aid lobby which long has been active and, increasingly, from old and new private school constituents who have a variety of vested interests. Included in the latter group are people who oppose racial integration in public schools and see the private school as a haven of homogeneity, radicals who improvise pseudo-educational operations to teach revolution, as well as sponsors of a variety of types of innovation in educational enterprises, many of which aim to serve culturally disadvantaged children and youths. Strange as the bedfellows are, they form a common chorus so strident and insistent that it strikes fear in the hearts of many state legislators and influences the votes of national congressmen as well as Presidential policies.

The appeals are familiar: patrons of nonpublic schools pay taxes and, thus, are entitled to a share of the revenues; private and parochial schools render needed public services, provide competitive alternatives to public schools that prod improvements in the latter, as well as insure a diversity in education that is healthy for the society. An additional current argument is that nonpublic schools are financially hard-pressed and would, if forced to close, increase the burden on public schools—at even greater costs than the amounts being requested for subsidies to private and parochial schools.

The current crusade for public support of nonpublic schools rides a crest of disenchantment with the public schools by contrasting segments of the society. The poor, racial, and ethnic minorities, as well as the rich who have traditionally shunned sending their children to the "common" schools, are found among the critics. A popular argument runs that the public schools are so tradition-bound, so inhibited by uniform standards (e. g., for teacher certification, accreditation, state prescribed curricula, school populations, and operating procedures), as well as so enmeshed in bureaucratic red tape, that they cannot respond to the unique needs of students who differ from the white, middle-class, Protestant majority. It is argued that only nonpublic schools, free from public school standards and accountability, can be innovative. As confidence in public schools is undermined, nonpublic institutions are eulogized as the perfect remedy—with public support and without its controls, of course.

Other factors that contribute to the illusion that the nonpublic school is the answer to educational prayers are some of the innovations currently being tried out to improve public schools. Performance contracts are being cited as evidence that private firms must be employed to accomplish what public school teachers cannot. Voucher plans are being touted as an escape route from the public school. Provisions of the Elementary and Secondary Education Act of the Federal Congress, which allocates funds for certain kinds of instructional resources and staff improvement services to nonpublic schools, are seen as the first crack in the wall of separation between church and state and public and nonpublic schools that can lead to more abundant support. Proposals by the President for revenue sharing that would provide funds for nonpublic as well as public schools are seen as the *coup de grace* to resistance.

While the crusade gains momentum, a wall of silence seems to hide some of the great social issues that are involved. Whether given proposals for public support for nonpublic schools are constitutional is a matter for the Supreme Court to decide. Within the confines of legal procedures, policies adopted bear heavily upon the objectives of the nation. Flat grants to nonpublic schools promise to promote a dual school system: private schools for the well-to-do who can pay the tuition difference between the grant and school costs, and public schools for the poor. They will tend to counteract the national commitment to school integration and to equality of educational opportunity. Voucher and purchase of service plans will have the same effect unless protected by safeguards that require nonpublic schools to admit and retain all students on an equal basis, without discrimination based on race, color, creed, physical and mental abilities, as well as on behavioral manifestations. Chaos will result if nonpublic schools that receive public funds are not required to adhere to standards set for the public schools.

A case can be made for utilizing the resources of nonpublic schools to accomplish public educational objectives. If voucher or purchase of services plans can bring benefits to students who need help most—the poor, the alienated, and the culturally deprived, rather than the rich and the religious—they should be utilized. It is a mistake to believe, however, that only schools without accountability to the public at large can innovate to meet student needs—public schools are the pioneers for change, not the private and parochial schools, as is being alleged. Nor is it sound to suggest that costs of education will be less in nonpublic schools. A dual school system, which is what many proposals would create, is always more expensive to maintain. Furthermore, public schools as well as nonpublic institutions are in financial distress. Many currently are not renewing contracts of as high as 10 percent of their teaching forces, not because of reductions in enrollments, but for lack of funds. Any plan to use the resources of nonpublic schools for public purposes should augment and strengthen, rather than weaken, public schools.

THE JOURNAL of EDUCATIONAL RESEARCH—October 1971

MADISON, WISCONSIN

Executive Editors

WILSON B. THIEDE, Professor of Curriculum and Instruction, The University of Wisconsin, Madison, Chairman.

WAYNE OTTO, Professor of Curriculum and Instruction, The University of Wisconsin, Madison.

L. JOSEPH LINS, Director of Research, Wisconsin Coordinating Committee for Higher Education, Madison, Wisconsin.

Consulting Editors

Terms Expire December 31, 1971

LAWRENCE P. BLUM, Professor of Educational Psychology, University of Wisconsin—Milwaukee.

DOUGLAS R. GROSS, Arizona State University, Tempe.

WILLIAM B. McCOLLY, Professor of English, University of South Carolina, Columbia.

WILLARD E. NORTH, Director of Research, Central Missouri State College, Warrensburg. Field News Editor.

JOHN R. PALMER, Professor of Curriculum and Instruction and History, The University of Wisconsin, Madison.

JOHN E. STECKLEIN, Bureau of Institutional Research, University of Minnesota, Minneapolis.

JULIA R. VANE, Professor of Psychology, Hofstra University, Hempstead, Long Island, New York.

Terms Expire December 31, 1972

NORMAN J. BOYAN, Dean, Graduate School of Education, University of California, Santa Barbara.

ROBERT D. BOYD, Associate Professor of Curriculum and Instruction, University of Wisconsin, Madison.

ROBERT CLASEN, Instructional Research Lab., University of Wisconsin, Madison. Book Review Editor.

CARROLL L. GONZO, Assistant Professor of Music Education, University of Illinois, Urbana.

BOB W. JERROLDS, Assistant Professor of Education, University of Georgia, Athens.

KARL KOENKE, Assistant Professor, College of Education, University of Illinois, Urbana.

HAROLD E. MITZEL, Assistant Dean for Research, College of Education, Pennsylvania State University, University Park.

Terms Expire December 31, 1973

ROGER W. BARDWELL, Superintendent of Schools, Ithaca City School District, New York.

DOLORES DURKIN, Professor of Elementary Education, University of Illinois, Urbana.

OLIVER L. HURLEY, Associate Professor of Special Education, Yeshiva University, New York.

GEORGE S. JURATA, Coordinator of Reading, Jefferson County Schools, Colorado.

JAMES D. RATHS, Professor of Education and Director, Bureau of Educational Research and Field Services, University of Maryland, College Park.

S. JAY SAMUELS, Associate Professor of Educational Psychology, University of Minnesota, Minneapolis.

NORMAN R. STEWART, Professor, Department of Counseling, Personnel Services and Educational Psychology, Michigan State University, East Lansing.

Editors Emeriti

HARRY J. BAKER, 2241 Q Via Puerta, Laguna Hills, California 92653.

FOWLER D. BROOKS, 724 N. 8th St., Neodesha, Kansas 66757.

HARL R. DOUGLAS, 4185 Pinion, Boulder, Colorado 80302.

EARL HUDELSON, 891 South Walnut St., Box 323, Anaheim, California 92802.

A. M. JORDAN, Professor Emeritus, School of Education, University of North Carolina, Chapel Hill, North Carolina 27514.

ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California 90027.

WILLIAM A. McCALL, 501 Alminar Avenue, Coral Gables, Florida 33146.

JESSE B. SEARS, 40 Tevis Place, Palo Alto, California 94301.

51st Year of Publication

The Journal of Educational Research

Dedicated to the Study of Education

VOLUME 65 OCTOBER 1971 NUMBER 2

ARTICLES

Student Needs, College Expectations, and "Reality" Perceptions—*Edwin L. Herr* 51

The i.t.a. Transition Problem—A Comparative Study—*Bernard J. Shapiro; Robert E. Willford; Phyllis P. Shapiro* 57

Focused and Unfocused Feedback and Self-Perception—*Agnes G. Rezler; Alexander S. Anderson* 61

Differences in Personality Traits Among Most Preferred and Least Preferred Students in Grades 10, 11, and 12—*Norman J. Bauer* 65

Relationship of Study Habits and School Attitudes to Achievement in Mathematics and Reading—*Florence P. Shepps; R. Ronald Shepps* 71

The Use in Writing of Textbook Methods of Paragraph Development—*Richard A. Meade; W. Geiger Ellis* 74

Evaluation of an Innovative Educational Device—A Case Study—*David P. Yens* 77

Open-Closed Mindedness and an Inquiry-Oriented Social Studies Methods Course—*James P. Shaver; Hyrum E. Richards* 85

DEPARTMENTS

Among Those Present 96

Field News—*Willard E. North* 94

Forthcoming Articles 73

It Says in The JER 50

Research Briefs—*Robert J. Thornton* 64

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $10.00 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies $1.50. Second class postage paid at Madison, Wisconsin.

Copyright, 1971 by Dembar Educational Research Services, Inc.

Address all correspondence care of DERS, Box 1605, Madison, Wisconsin 53701.

Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, President. Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistant to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. SANDRA BENTHEIMER LEWIS, Supervisor of Editorial Services.

It Says in The JER...

STUDENT NEEDS and expectations of college were related across sixteen of thirty person situation dimensions in Herr's study. He also found that student expectations of college and perceptions of college realities were related in thirty possible cases. Sex of the student influenced the magnitude of the latter correlation (p. 51).

TO READ or not to read is *not* the question—the instructional strategies in transferring from an initial teaching alphabet to traditional orthography are in question. Shapiro, Willford, and Shapiro found that the more structured transition strategies appeared to be the more effective ones (p. 57).

IMPROVING SELF-AWARENESS was the purpose of a workshop for health educators which used focused (video tape stopped and viewer attention directed to specific cues) and unfocused (viewers watch tape without interruption) video tape review. Rezler and Anderson found a need for repeated, focused reviews instead of single self-confrontations (p. 61).

SOCIAL STANDARDS, social skills, and antisocial tendencies were just some of the personality traits found to be significantly different between most and least preferred adolescents. Bauer also found differences on traits such as family relations, personal freedom, and self-reliance (p. 65).

DO STUDY HABITS and school attitudes bear any relationship to achievement in sixth grade math and reading? Shepps and Shepps, using The Survey of Study Habits and Attitudes (SSHA), found that in the case of boys, the total SSHA score and the attitudes subtest predicted reading achievement. In the case of girls, the attitudes subtest predicted math achievement (p. 71).

TEXTBOOKS SUGGEST various methods of paragraph development. Meade and Ellis found that over half the paragraphs in contemporary expository writing that they studied were not developed by textbook methods. However, three non-textbook methods of development were clearly identified: opposition, question, and additional comment (p. 74).

STUDENTS LIKE getting immediate test scores and a mechanical test-scoring machine helps make this possible. Yens studied students in grades 4, 7, and 10 and found that the machine-using groups showed no superiority over groups using a quiz only approach (p. 77).

RESULTS IN the Shaver and Richards study of education students indicate little relationship between open-closed mindedness and reactions to an inquiry-oriented social studies method course. There were a scattering of significant correlations between F-, D-, and R-scores and student ratings of instructors, but no pattern to show a definite consistent relationship betweeen Ss' open-closed mindedness and their evaluation of the instructor (p. 85).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 65, Number 2, October 1971)

Student Needs, College Expectations, and "Reality" Perceptions[1]

EDWIN L. HERR, The Pennsylvania State University

ABSTRACT

The purpose of this study was to test four hypotheses concerned with relationships between student needs, college expectations, perceptions of college realities, and four groups of student outcome: status quo, changed major, transfer, and withdrawal. Pair-wise Pearson product-moment correlations were obtained between like scales on the Stern Activities Index and two administrations of the College Characteristics Index, prior to college matriculation and at the conclusion of the freshman year. Student needs and expectations of college were related across sixteen of thirty person-situation dimensions but the magnitude and direction of correlations was strongly influenced by sex of respondent. Student expectations of college and subsequent perceptions of college realities were significantly correlated in thirty possible cases. The S's sex influenced the magnitude of the correlations obtained between college expectation and college realities.

GRADUATES OF American high schools are entering institutions of higher learning in rapidly increasing numbers. In 1964, there were 4,800,000 students enrolled in higher educational institutions, a figure double that of a decade earlier (9). By 1980, it is forecast that there will be 17,051,000 Americans between 18 and 21 years of age of whom approximately 60 percent will be attending institutions of higher learning; 3,000,000 students will be entering freshmen (1).

As a result of rapid technological change, expanded educational and occupational horizons among all social classes, and greatly increased financial assistance available to those who desire further training, it can be expected that the college choice process with its attendant subdecisions will become evermore complex. Even now, efforts to facilitate the transition from high school to college are less than satisfactory. One spokesman for higher education (1) maintains that, *"because of our collective failures in both secondary and higher education, we have come close to making access to our colleges and universities a shambles."* Available evidence about student attrition in college tends to confirm the fundamental validity of this observation as well as of others (2) more directly focused on needed modifications in guidance and counseling approaches to the facilitation of student decisions about college selection. At the present time, as few as three out of five students in college stay continuously to graduation (10). *"In general, student dropouts arise largely from failure to meet psychological, sociological, or economic demands rather than the strictly academic demands of college life"* (8). Among the factors apparently influencing this phenomenon is an inappropriate original choice of institution (4).

With these broad descriptions of the high school to college transition as reference points, the present pilot study (3) was promulgated to examine the following hypotheses:

1. Student expectations of college will not be associated with student needs.
2. Student expectations of college will not be congruent with their perceptions of the actual college environment.
3. Lack of congruence between expectations of college and perceptions of college will not be associated with plans for withdrawal, transfer, or change of course.
4. Lack of congruence between students' needs and perception of college will not be associated with plans for withdrawal, transfer, or change of course.

Sample

The total sample for the initial data collection consisted of 231 seniors from seven senior high schools in western New York State. The criterion for selection of the sample was receipt of an acceptance from a college which the student subsequently planned to attend. The schools represented metropolitan comprehensive, surburban college-oriented, surburban heterogeneous, semi-rural, and rural high schools. The communities in which these schools were located reflect a wide range of socioeconomic strata although, as might be expected, the student sample skewed toward the middle and upper socioeconomic cluster of the range. The general educational level of the families of these students was also skewed toward the upper end of the educational range. Most of the families placed a great deal of importance on their child attending college.

A majority of students in the sample (68%) had made a choice of major field of study, and most of them had made this choice within the last 2 years. Further, 48 percent stated that they had no difficulty in making the choice.

Methodology and Instrumentation

The students were tested over a 2-day period during the months of April and May 1965. Of the initial sample, 231 completed the testing and provided usable data for the study. During April and May 1966, at the conclusion of the freshman year in college, the 231 students were again contacted. After three follow-up attempts, 158 students responded with usable data.

The test battery administered to the sample consisted of two administrations of the College Characteristics Index (CCI), the Activities Index (AI), and a 2-part, locally-devised questionnaire designed to gather biographical information.

The CCI and the AI are each instruments consisting of thirty scales of ten items or a total of three hundred statements. Twenty-two of the scales are bipolar with a high scale score indicating a need or press for the first adjective. The first instrument (CCI) is designed to assess the environmental "press" of the institution examined: e.g., the assumed behavioral or psychological demands which characterize the institution and which facilitate or impede individual expression of needs. The AI provides a measure of individual needs. The two insrtuments are parallel in format, scale items, and scoring in order to make comparable analysis of person-situation parameters possible.

Following are the scale labels and their abbreviations as well as the assignment to press clusters as used in Tables 1–4 (5).

Intellectual Need-Press
Achievement (ach)
Conjunctivity-Disjunctivity (cnj-dsj)
Counteraction-Inferiority Avoidance (ctr-inf)
Ego Achievement (e/a)
Energy-Passivity (enj-pas)
Humanism (hum)
Objectivity-Projectivity (obj-pro)
Reflectiveness (ref)
Science (sci)
Understanding (und)

Dependency Need-Press
Abasement-Assurance (aba-ass)
Adaptability-Deference (ada-def)
Affiliation-Rejection (aff-rej)
Deference-Restiveness (def-rst)
Nurturance-Rejection (nur-rej)
Order-Disorder (ord-dso)
Supplication-Autonomy (sup-aut)

Emotional Expression Need-Press
Aggression-Blame Avoidance (agg-bla)
Change-Sameness (cha-sam)
Dominance-Tolerance (dom-tol)
Emotionality-Placidity (emo-plc)
Exhibitionism-Inferiority Avoidance (exh-inf)
Fantasied Achievement (f/a)
Impulsiveness-Deliberation (imp-del)
Narcissism (nar)
Play-Work (ply-wrk)
Practicalness-Impracticalness (pra-ipr)
Harm Avoidance-Risktaking (har-rsk)
Sensuality-Puritanism (sen-pur)
Sexuality-Prudishness (sex-pru)

The CCI was first administered to students prior to actual matriculation at college with instructions to complete it in terms of what they would expect to find at the institution where they had been accepted and intended to matriculate. The CCI was administered a second time after students had been enrolled approximately one year in the college to which they had been accepted at the time of the first testing in order to determine their perceptions of the realities of the college environment. These data in combination were used to examine relationships between needs and expectations of college, college expectations and perceptions of actual college experiences, as well as the effects of congruity between college expectations and reality as manifested in future collegiate plans. Student responses to the various instruments administered were stratified further on the basis of sex, grade point average (GPA), and high school attended.

This paper discusses only the results obtained when four hypotheses from the larger study were tested by pair-wise Pearson product-moment correlations between each pair of scales carrying the same label on the AI (needs) and the CCI (press) or between the pre-CCI (college expectations) and the post-CCI (college realities). The AI scale reliabilities (K–R Formula 20) range from .40 to .88 with average reliability = .69. Item discrimination values, reflecting internal consistency, range from .27 to .81 and average .57. CCI scale reliabilities range from .34 to .81 with average reliability = .65. Item discrimination values range from .43 to .66 with the average = .52. Stern has noted that the average reliability is close to the maximum for scales of ten items (7). The relatively high discrimination values indicate that the content of the scales is homogenous (6).

Limitations

The limitations of this research are as follows: 1. The instruments used are parallel assessments of the person-situation parameters considered to be important in the frame of reference from which this study was generated. Although the theoretical assumptions upon which the instrument construction is based can be readily documented, the specific item construction cannot be as easily defended. 2. The sample size, although sufficiently large at the outset, became dissipated in follow-up. 3. The degree to which the relationships observed in this study were conscious or unconscious inputs to student choice-making are unanswered.

Findings and Conclusions

Hypothesis 1: *Student expectations of college will not be associated with student needs.*

When Ss were combined into one group, sixteen significant correlations were obtained between personal needs and expectations of college (12, $p < .01$; 4, $p < .05$). For data interpretation, the thirty scales were divided into three dimensions of college climate—intellectual, dependency, and emotional expression (5). Significant correlations were obtained between five scales designated as intellectual; four, dependency; and seven, emotional expression. The way students described their personal needs was significantly related to that which they expected to characterize the college environment on sixteen of thirty possible need-press dimensions.

When needs and college expectations were separately correlated for males and females, a smaller number of significant correlations were obtained: seven for males and twelve for females. Five of the combined scales yielded significant correlations for both sexes, although on the scales identified as assessing emotional expression need-press females ordered their expectations of college with greater correspondence to their self-descriptions than did males. Neither GPA nor sending school of the respondents appeared to be as powerful an influence on expectations of college as did sex.

As can be inferred from the differences in magnitude and, in some instances, directions of the correlations as well as by inspection of scale means and standard deviations which are not reported here, it is clear that the student's sex does discriminate personal needs and expectations of college. As a result of the complex factors identified here Hypothesis 1 can not be accepted.

Hypothesis 2: *Student expectations of college will not be congruent with their perceptions of the actual college environment.*

When Ss were combined, thirty significant correlations ($p < .01$) were obtained between

Table 1.—Correlations Between Student Needs (AI) and College Expectations (Pre-CCI) for all Ss, for Males, and for Females

Intellectual Need-Press

	Comb	M	F
ach	.088	.097	.085
cnj-dsj	.092	.147	.017
ctr-inf	.087	.069	.135
e/a	.284**	.242*	.333**
enj-pas	.072	.035	.075
hum	.339**	.315**	.258**
obj-pro	.020	.029	-.032
ref	.220**	.194	.213*
sci	.248**	.303**	.172
und	.289**	.269**	.318**

Dependency Need-Press

	Comb	M	F
aba-ass	.031	.101	.029
ada-def	.138*	.090	.167
aff-rej	.122	.197*	-.002
dfr-rst	.235**	.163	.231*
nur-rej	.291**	.248*	.213*
ord-dso	.140*	.101	.088
sup-aut	.126	.040	.152

Emotional Expression Need-Press

	Comb	M	F
agg-bla	.244**	.089	.258**
cha-sam	.277**	.175	.400**
dom-tol	.098	-.054	.145
emo-plc	.181**	.145	.293**
exh-inf	.090	.049	.088
f/a	.108	.037	.168
imp-del	.135	.146	.131
nar	.153*	.107	.162
ply-wrk	.214**	.213	.205*
pra-ipa	.067	.044	.089
har-rsk	.282**	.265**	.213*
sen-pur	.063	-.080	.120
sex-pru	.171*	.130	.265**

Comb * .05 = .138
** .01 = .181
DF 223

Male * .05 = .195
** .01 = .254
DF 106

Female * .05 = .195
** .01 = .254
DF 117

Table 2.—Correlations Between Student Expectations of College (Pre-CCI) and Student Perceptions of College Realities (Post-CCI) for all Ss, for Males, and for Females

Intellectual Need-Press	Comb	M	F
ach	.457**	.468**	.425**
cnj-dsj	.305**	.179	.419**
ctr-inf	.291**	.159	.406**
e/a	.329**	.337**	.362**
enj-pas	.471**	.472**	.499**
hum	.577**	.613**	.618**
obj-pro	.256**	.248*	.335**
ref	.534**	.598**	.501**
sci	.512**	.546**	.475**
und	.352**	.412**	.334**

Dependency Need-Press	Comb	M	F
aba-ass	.228**	.237*	.280*
ada-def	.229**	.168	.268*
aff-rej	.340**	.416**	.240*
dfr-rst	.333**	.169	.441**
nur-rej	.494**	.372**	.538**
ord-dso	.383**	.293**	.462**
sup-aut	.414**	.311**	.508**

Emotional Expression Need-Press	Comb	M	F
agg-bla	.488**	.355**	.500**
cha-sam	.405**	.440**	.403**
dom-tol	.366**	.326**	.374**
emo-plc	.347**	.450**	.229
exh-inf	.304**	.387**	.317**
f/a	.421**	.560**	.330**
imp-del	.342**	.254*	.372**
nar	.490**	.326**	.618**
ply-wrk	.376**	.421**	.337**
pra-ipr	.435**	.375**	.446**
har-rsk	.382**	.012	.431**
sen-pur	.536**	.542**	.542**
sex-pru	.481**	.474**	.490**

Comb * .05 = .159
** .01 = .208
DF 153

M * .05 = .232
** .01 = .302
DF 78

F * .05 = .232
** .01 = .302
DF 75

scales assessing expectations of college (pre-CCI) and the realities of college (post-CCI). Hence, the way students ordered their expectations of college were positively and consistently related to that which they found to exist in the college environment to which they were exposed.

For males twenty-five of thirty possible correlations were found to be significant and for females twenty-nine of thirty. Again, the differential magnitude of correlations from scale to scale compared by sex indicates that the expectations of college and the college realities reported were sufficiently unique to males and to females to warrant more refined investigation even though the significance of the correlations indicates the general congruence of college expectations and college realities. Hypothesis 2 was rejected.

Hypothesis 3: *Lack of congruence between expectations of college and perceptions of college will not be associated with plans for withdrawal, transfer, or change of major.*

To test Hypothesis 3, students were classified into four groups: Status Quo (SQ)—those who after the freshman year intended to stay in the same institution and the same major field of study; Changed Major (CM)—those who intended to remain in the same institution but change major field of study; Transfer (T)—those who intended to transfer to another collegiate institution; and Withdrawals (W)—those who intended for the immediate future to completely withdraw from college. Because of attenuation in the follow-up sample, the latter two groups were too small for meaningful interpretation. However, the existing data suggested rather gross differences between what press was expected to characterize the college environment and what was actually found. Certainly those who changed major fields of study reported less congruence between expectations of college and perceptions of college reality, at least in a global sense, than did those who remained in the original choice of institution and major field of study. Hypothesis 3 was tentatively rejected.

Hypothesis 4. *Lack of congruence between student needs and perceptions of college will not be associated with plans for withdrawal, transfer, or changes of major.*

The same classification scheme was used to test

Table 3.—Correlations Between Expectations of College and Perceptions of College for SQ, CM, T, and W Students

Intellectual Need-Press

	SQ	CM	T	W
ach	.507**	.331	.228	.885*
cnj-dsj	.307**	.080	.851*	.945*
ctr-inf	.277**	.213	.244	.767
e/a	.314**	.398	.688	.148
enj-pas	.464**	.622**	.170	.408
hum	.620**	.327	.565	.799
obj-pro	.229**	.449*	.321	-.222
ref	.570**	.690**	.318	.479
sci	.515**	.628**	.611	.866
und	.402**	.121	.470	.395

Dependency Need-Press

	SQ	CM	T	W
aba-ass	.110	.521*	.938**	.044
ada-dfs	.243*	.103	.486	.629
aff-rej	.362**	.528*	-.638	.400
dfr-rst	.374**	-.095	-.679	-.310
nur-rej	.502**	.455*	.583	-.139
ord-dso	.372**	.501*	.166	-.238
sup-aut	.465**	.250	-.217	.656

Emotional Expression Need-Press

	SQ	CM	T	W
agg-bla	.469**	.648**	.425	.702
cha-sam	.382**	.562**	.768	-.240
dom-tol	.397**	.056	.692	-.129
emo-plc	.201*	.681**	.772	.319
exh-inf	.357**	.234	.490	-.704
f/a	.429**	.489*	.650	.634
imp-del	.379**	.054	.213	.139
nar	.517**	.423*	.731	.714
ply-wrk	.389**	.376	.127	-.201
pra-ipr	.385**	.603**	.584	.919*
har-rsk	.435**	.202	-.228	.584
sen-pur	.545**	.284	.493	.957*
sex-pru	.506**	.342	.812*	.075

SQ	CM	T	W
* .05 = .195	* .05 = .423	* .05 = .811	* .05 = .878
** .01 = .254	** .01 = .537	** .01 = .917	** .01 = .959
DF 121	DF 22	DF 6	DF 5

Hypothesis 4 as was used to test Hypothesis 3. The same attenuation occurred in the groups making the data on T and W outcomes virtually useless. Nine significant correlations were obtained between personal needs and perceptions of college reality for the SQ group; four for the CM group; none for the T's; and two for the group intending to withdraw. An interesting trend in the data concerns the increasing numbers of negative correlations, even though nonsignificant, which can be observed as data from the SQ group is contrasted with the other groups. This phenomenon suggests that students in groups drastically altering their college plans perceived little or no relationship between their personal needs and that which the college rewarded or that the college environment rewarded behavior inversely related to student preferred or actual behavior. Hypothesis 4 was accepted.

Discussion

The data reported in this study suggest that the hypotheses examined oversimplify some of the complex influences which attend the college choice. Student descriptions (needs) and the expectations of college which they hold, and which supposedly enter into their choice of college, are related among many dimensions of the person-situation continuum. The sex of the respondent is significantly related to self-description and environmental perception. However, this study does not answer whether the relationship of needs or self-descriptions and college environmental dimensions are conscious manifestations in the above process. When student self-descriptions and expectations of college are examined there are indications that students do not expect needs and related press to be of the same magnitude. Apparently, on some self-situation dimensions, the college environment is expected to exert a drawing out or expanding effect upon the individual while on other dimensions the college environment is expected to restrain or not reward particular individual behavior.

The findings, then, raise questions such as: What is the magnitude, if any, of the gap between what students expect in college and what they find to lend stimulus to their personal efforts to achieve their original goals? Does college represent some real-ideal relationship for students to enable them to grasp for a vision of what they can become or what they must become to succeed and realize satisfaction in the vagaries of the many faceted college environment? Answers to these questions represent the needed thrust for further research in the areas of the non-intellec-

Table 4.—Correlations Between Student Needs and Perceptions of College for SQ, CM, T, and W Students.

Intellectual Need-Press

	SQ	CM	T	W
ach	.227*	.372	-.193	.411
cnj-dsj	.110	.197	.520	.350
ctr-inf	.187	.170	.523	-.049
e/a	.128	.092	-.298	-.042
enj-pas	.032	.146	-.779	.667
hum	.420**	-.016	-.135	.630
obj-pro	.055	-.159	.245	.066
ref	.157	.475*	-.295	.179
sci	.421**	.490*	.513	.592
und	.314**	.373	.163	.000

Dependency Need-Press

	SO	CM	T	W
aba-ass	-.062	.033	-.243	.639
ada-dfs	.040	-.021	-.009	-.338
aff-rej	.078	-.110	-.555	-.286
dfr-rst	.069	.454*	-.286	.029
nur-rej	.197*	.265	.047	.200
ord-dso	.044	.380	.190	.082
sup-aut	.213*	-.037	-.744	.924*

Emotional Expression Need-Press

	SQ	CM	T	W
agg-bla	.317**	.367	.448	.503
cha-sam	.040	.078	-.108	-.301
dom-tol	.081	-.005	.317	.486
emo-plc	.054	.275	.058	-.317
exh-inf	.047	-.052	-.264	-.936**
f/a	.161	.011	.287	.285
imp-del	.204*	.062	-.232	-.447
nar	.272**	.221	.420	.000
ply-wrk	-.040	-.195	-.046	-.012
pra-ipr	.134	.632**	.406	-.408
har-rsk	.042	.258	-.197	.450
sen-pur	.032	.102	-.125	-.018
sex-pru	.097	.346	-.000	-.482

SQ	CM	T	W
* .05 = .195	* .05 = .423	* .05= .811	* .05 = .878
** .01 = .254	** .01 = .537	** .01 = .917	** .01 = .959
DF 120	DF 22	DF 6	DF 5

tive comparisons of self-situation parameters as these pertain to educational planning. It is true that the disparities between college expectations and college demands are real and that the increased availability of information in these areas effectively used by counselors could contribute much important input to the decision-processing of young people contemplating college.

FOOTNOTE

1. This study was funded by a grant from the College Entrance Examination Board.

REFERENCES

1. Crossland, F. E., "Politics and Policies in College Admission," *Phi Delta Kappan,* 46:299–302, 1965.
2. Herr, E. L., "Guidance of the College-Bound," *Personnel and Guidance Journal,* 45:47–52, 1966.
3. Herr, E. L.; Moore, G. D., A Pilot Study of College Expectations and "Reality" Perceptions, Cooperative Research Project, College Entrance Examination Board and State University of New York at Buffalo, 1967.
4. Iffert, R. E., *Retention and Withdrawal of College Students,* U. S. Government Printing Office, Washington, D. C., 1958.
5. Stern, George G., "Continuity and Contrast in the Transition from High School to College," in *Orientation to College Learning—A Reappraisal,* American Council on Education, Washington, D. C., 1961.
6. Stern, G. G., "Environments for Learning" in Sanford, R. Nevitt (ed.), *The American College: A Psychological and Social Interpretation of the Higher Learning,* John Wiley, New York, 1962.
7. Stern, G. G., "The Measurement of Psychological Characteristics of Students and Learning Environments," in Messick, S. J.; Ross, J. (eds.), *Measurements in Personality and Cognition,* John Wiley, New York, 1962.
8. Summerskill, J., "Dropouts from College," in Sanford, R. Nevitt (ed.), *College and Character,* John Wiley and Sons, Inc., New York, 1964.
9. U. S. Office of Education, *Digest of Educational Statistics,* U. S. Government Printing Office, Washington, D. C., 1964.
10. Wise, W. M., "The American Student and His College," in *From High School to College: Readings for Counselors,* College Entrance Examination Board, Princeton, New Jersey, 1965.

THE simple realization that there are other points of view is the beginning of wisdom. Understanding what they are is a great step. The final test is understanding why they are held.
—*C. M. Campbell*

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 65, Number 2, October 1971)

The i.t.a. Transition Problem—A Comparative Study

BERNARD J. SHAPIRO
Boston University

ROBERT E. WILLFORD
Radiant Corporation

PHYLLIS P. SHAPIRO
Emmanuel College

ABSTRACT

The purpose of this study was to investigate the measured effectiveness of three instructional strategies as applied to the reading, spelling, and language achievement of children undergoing transition from i.t.a. to T.O. Twelve second grade classes were randomly assigned to one of the three strategies. The findings indicate that although at the end of the transition period all groups had made statistically significant gains over their pre-transition level, the more structured transition strategies appeared to be the more effective ones. In any case, it appears that the transition phase need not be a time of regression or status quo in terms of reading achievement.

AS A RESULT of various i.t.a. projects (e.g., 3), attention has become focused on that aspect of the i.t.a. system known as "transition." Although many have suggested that transition from the initial teaching alphabet (i.t.a.) to the traditional orthography (T.O.) was an automatic process, the research data (e.g., 1, 2, 3) has tended to show that there were considerable difficulties involved. Our own observation of i.t.a. project classes indicated that:

1. *The assigned level of post-transition reading material in T.O. was typically below the readability level of the last book in the i.t.a. program.*
2. *Teachers found it necessary to do considerable post-transition diagnostic teaching in order to enable the child to handle T.O. materials adequately.*
3. *Children continued to use i.t.a. symbols in their written work.*

Thus, it would seem that transition was not an incidental or automatic process; rather it appeared to be a complex stage of learning that warranted both comprehensive instruction and more detailed research.

The purpose of this study was to investigate the measured effectiveness of three instructional strategies as applied to the task of mediating reading, spelling, and language achievement of children undergoing transition from i.t.a. to T.O.

Instructional Strategies

Instructional strategy I (referred to in this study as the ETR strategy) included the procedures, methods, and materials described in the transitional phase of the i.t.a. reading program, entitled the Early to Read Series, plus an undefined post-transitional program. The Early to Read Series is published by Initial Teaching Alphabet Publications, Inc., of New York. The transitional phase of this series is limited to portions of Books 6 and 7 (First Edition), the accompanying workbook, and the teacher's manual. The instructional strategies used in this program to elicit transitional decoding and encoding behavior are described in the teacher's manual and inferred from the readers and workbook. The procedure used in the expository reading materials to introduce i.t.a. configurational and traditional respelling changes involves a combination of dissolution of the unique features of the i.t.a. augmentations and the injection of non-prompted traditional respellings of certain phonemes. The method followed in the transitional workbook involves eliciting certain marking and respelling responses in the presence of word and pictorial stimuli. The formal transition phase of this program involved approximately 55 days of instructional time to complete. At the termination of this phase of the strategy, the Ss were introduced to various post-transitional T.O. reading programs.

Instructional Strategy II (referred to as the GCRP study) involved the procedures, methods, and materials included in the Greater Cleveland Reading Program—GCRP/i.t.a. Language Arts Transition Program. This is an experimental program designed specifically for this study. The program consists of eighty-seven instructional units. Each unit is designed to elicit specific responses identified as crucial to the demonstration of adequate transitional behavior. The strategies used in this program are related to the tasks of (a) bringing learned verbal and written responses under the control of new stimuli, (b) eliciting additional differentiated responses to discriminative stimuli, and (c) eliciting new verbal, intraverbal, and written responses in the presence of stimuli of increasing complexity. The program consists of a teacher's manual, a student workbook, supplemental work pages, wall charts, flashcards, and two transitional readers.

Instructional Strategy III (referred to as the "no transition" or "control" strategy) involved Ss making immediate transfer from instruction and reading in the i.t.a. medium to instruction and reading in T.O. without exposure to any instruction of transition elements. The change took place at the beginning of the experimental period when the Ss had reached the transition phase of Book 6 of the first edition of the Early To Read series.

Methodology

Sample. At the beginning of the school year, the investigators selected twelve second grade classes in which all of the students had received their basic i.t.a. instruction through the Early To Read Series but had not yet progressed beyond the series' introduction of the first transition elements. Four of the classes were randomly assigned to each of the instructional strategies. Prior to the start of school, the teachers were notified of the assignments and arrangements were made for in-service workshops designed to familiarize teachers with the program they were to use. During the study itself, observers visited each of the classes twice weekly to insure that the guidelines for each of the instructional strategies were being followed.

The Testing Program. In addition to the Lorge Thorndike Intelligence Test (Level 2, Form B), two pretests were administered. Forms X (i.t.a. version) and W (T.O. version) of the Stanford Achievement Test's Primary II Battery were used so as to provide initial achievement levels in both i.t.a. and T.O. Form Y (T.O. version) of the same battery was given to all Ss when those using the ETR materials had completed their formal transition activities. The final posttest (Form X—T.O. version) was administered at the end of the treatment period, i.e., when Ss in the GCRP group had completed the eighty-five lessons.

Findings. The mean IQ's of the GCRP, ETR, and control groups were 109.92, 108.24, and 108.92 respectively, and an analysis of variance ($F_{2, 24} = .18$) indicated that these differences were not statistically significant at the 5 percent level.

Since intact classes rather than individual Ss had been assigned to the treatment groups, it was necessary in the consideration of the achievement test scores to insure against differences between the groups in their pre-transition level of i.t.a. reading achievement. Therefore, in considering the final posttest measures, the analysis of covariance was used with the i.t.a. pretest as the covariate. The results of these covariance analyses are presented in Table 1.

As indicated by the F ratios in Table 1, statistically significant differences between the groups were observed in all the Stanford subtests with the exception of Paragraph Meaning. In investigating these general results more specifically, the Duncan Multiple Range Test indicated that:

1. *In Word Meaning and Word Study Skills, the GCRP mean was significantly greater than those of both the ETR and control groups.*
2. *The mean Spelling scores of both the GCRP and the control groups were greater than that of the ETR group.*
3. *The Language score of the GCRP group was significantly greater than that of the ETR group which in turn was significantly greater than that of the control group.*

In general, these findings indicate that when adjustment is made for initial level of i.t.a. reading achievement, alternate programs for i.t.a. transition instruction have direct influence on the immediate post-transition level of T.O. reading achievement. In terms of specific instructional strategies, the GCRP strategy seems, in general, to have been the most efficacious, although in the reading comprehension task (i.e., Paragraph Meaning), there were no statistically significant differences between any of the treatment groups.

In addition to the analysis of the final posttests, the present investigators wished to consider the general trend or pattern of achievement over the four administrations of the Stanford Battery. Thus, one-way analyses of variance between the four grade equivalent scores for each treatment group on each of the Stanford subtests were executed. The results of these analyses are presented in Table 2, and as can be seen from the table, the F ratios were in each case statistically significant at the 5 percent level. That is, within each treatment group and on each subtest, there were statistically significant changes in the measured level of achievement over the treatment period.

What kinds of trends are brought out by these differences? Two general patterns appear. First, for the Word Meaning, Spelling, and Word Study Skills tests, the pre-transition T.O. mean score is substantially lower than that of the pre-transition i.t.a. mean score, but this condition is followed by steady improvement for all groups. With the Word Meaning and Spelling tests, the final means are higher than the original i.t.a. levels, but with Word Study Skills, the post-

Table 1.—Analyses of Covariance: Post-Transition Scores—Criterion: SAT, Primary II, Form X (T.O.); Covariate: SAT, Primary II, Form X (i.t.a.)

SAT Subtest	Unadjusted Means			Adjusted Means			F
	GCRP	ETR	Control	GCRP	ETR	Control	
Word Meaning	25.52[a] (3.8)	21.40 (3.1)	21.56 (3.3)	23.81 (3.5)	22.11 (3.2)	21.40 (3.1)	7.12*
Paragraph Meaning	37.64 (3.2)	32.88 (3.0)	30.08 (2.8)	34.27 (3.0)	33.69 (3.0)	32.64 (3.0)	.24
Spelling	19.44 (3.6)	11.28 (2.8)	14.84 (3.2)	16.67 (3.4)	12.87 (3.0)	16.02 (3.3)	3.85*
Word Study Skills	50.40 (4.8)	43.00 (3.6)	41.56 (3.5)	49.63 (5.0)	43.26 (3.6)	42.07 (3.5)	2.97*
Language	47.20 (3.7)	41.80 (3.2)	34.84 (2.7)	45.85 (3.9)	42.49 (3.4)	35.49 (2.8)	5.88*

* Significant at the 5 percent level.

a. Main table entries are raw scores; figures in parentheses are the corresponding grade equivalent scores (students at grade placement 2.6).

transition means for the three groups were all below the level of their respective pre-transition i.t.a. scores. Second, for the Paragraph Meaning and Language tests, the T.O. pre-transition mean tended to be somewhat higher than the i.t.a. pre-transition mean, and there was steady improvement over time, the final post-transition scores being uniformly greater than those of the pre-transition measure.

Thus, in general, all groups at the end of the transition period had made gains over their pre-transition level of reading achievement. Prior to formal transition, the groups in this study did not demonstrate an ability to achieve at an equivalent level in i.t.a. and T.O. when the tasks involved Word Meaning, Spelling, and Word Study Skills. On the other hand, the reading skills measured by the Paragraph Meaning and Language subtests appear to be less influenced by immediate or "forced" transition.

Implications

Among the implications of the evidence presented are:

1. It is possible that much of the research evidence which has shown that the achievement level of i.t.a. children decreased during and following their transition experience may well be an indication of the quality of the transition program rather than an inherent transfer problem.
2. The fact that the i.t.a. children exhibited reasonable levels of achievement on the T.O. form of the pre-transition measure indicates considerable incidental T.O. learning without formal transition instruction. Given sufficient time, it is possible that the children's T.O. achievement might have been brought up to their i.t.a. pre-transition levels in all areas. There is, however, the question of both the time required and the comprehensiveness of the skills covered in such unstructured transition experiences.
3. The i.t.a. transition stage need not be a time when children either regress or remain at status quo in terms of their measured reading achievement. On the contrary, the transition stage is a time when children can expand their language

Table 2.—Mean Grade Scores Stanford Achievement Test, Primary II Battery Forms W, X, and Y

Group	Subtest	Form X(i.t.a.) GP 2.0	Form W(TO) GP 2.0	Form Y(TO) GP 2.4	Form X(TO) GP 2.6	F
GCRP	Word Meaning	3.08	2.79	3.47	4.01	5.49*
	Paragraph Meaning	2.53	2.70	3.36	3.59	5.13*
	Spelling	3.50	1.89	3.23	3.83	36.78*
	Word Study Skills	5.68	3.38	4.23	5.22	13.28*
	Language	3.37	2.70	3.50	4.06	9.53*
ETR	Word Meaning	2.33	2.10	2.75	3.31	14.59*
	Paragraph Meaning	2.08	1.98	2.71	3.10	15.09*
	Spelling	2.67	1.59	2.64	2.78	32.35*
	Word Study Skills	5.14	2.16	3.16	4.07	28.59*
	Language	2.12	2.33	2.96	3.26	18.13*
CONTROL	Word Meaning	2.22	2.12	2.61	3.16	20.11*
	Paragraph Meaning	1.94	1.97	2.68	2.82	35.05*
	Spelling	2.76	1.66	2.64	3.08	37.53*
	Word Study Skills	5.23	2.26	3.40	3.84	24.43*
	Language	2.12	2.45	2.70	3.11	19.95*

* Significant at the 5 percent level.

arts achievement in a continuous positive growth pattern.

REFERENCES

1. Downing, J., *The i.t.a. Symposium*, National Foundation for Educational Research in England and Wales, London, 1967.
2. Shapiro, B. J., *A Comparison of the Reading Achievement of i.t.a. and T. O. Groups in the First Grade. 1964–1965*, Educational Research Council of America, Cleveland, Ohio, 1966.
3. Shapiro, B. J., *A Longitudinal Study of i.t.a. A Comparison of the Reading Achievement of i.t.a. and T.O. Groups in the First Grade, 1965–1966*, Educational Research Council of America, Cleveland, Ohio, 1967.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 65, Number 2, October 1971)

Focused and Unfocused Feedback and Self-Perception

AGNES G. REZLER and ALEXANDER S. ANDERSON
University of Illinois, Chicago

ABSTRACT

The purpose of this study was to examine the effects of "focused" versus "unfocused" video tape review on self-perception. Focused review means that the video tape is stopped at selected places and the viewer's attention is directed to specific cues in order to alter his self-perception and subsequently his behavior. Unfocused review, on the other hand, means that viewers look at their taped performance without interruptions. The sample consisted of twenty-one health educators who attended a 1-week workshop to improve their self-awareness. Half of the sample was exposed to focused video tape review, the other half to unfocused review. Self-ratings were collected before and after the reviews to measure changes in self perceptions. The first finding of this study was that video tape reviews had to be focused for health educators to realize how rarely they engaged in reinforcing and idea seeking behaviors in small peer groups. The second finding pointed up the necessity of *repeated,* focused video tape reviews instead of single self-confrontations. Self-perceptions changed only temporarily under the impact of a single, focused video tape review.

IN MOST instances, there appears to be a gap between fact and perception regarding an individual's self-image. "To see oneself as others see us" has always been a challenge. The primary purpose of this article is to report some interesting data about the use of a relatively new technique, "self-confrontation," as a means of producing changes in individual's self-perception. Use of the technique involves viewing oneself on video tape in either a focused feedback or an unfocused feedback situation. Focused feedback implies that the video tape is stopped at selected places and the viewer's attention is directed to specific cues in order to alter his self-perception and subsequently, his behavior. Unfocused feedback, on the other hand, is defined as viewing oneself on video tape without interruptions and drawing one's own conclusions about what is being seen and heard.

Self-confrontation has been used with patients in individual and group therapy (2, 5, 6, 7) and in counselor and teacher training to provide accurate and specific feedback (4, 10). The concept of focused feedback was developed by Stoller (9) in the use of video tape in group therapy. Recently, Danet (3) made an attempt to submit the use of audio-visual feedback in group therapy to experimental study. The data suggested that unfocused feedback served to disrupt the group's functioning which supports Stoller's notions regarding the need for focused feedback.

The study reported herein addresses itself to the question: *Is self-confrontation a useful technique for producing changes in self-perception, or do viewers benefit from self-confrontation only if a supervisor, group leader, or therapist helps them focus on specific cues and behaviors and provides adequate opportunity for dealing with the material reviewed?*

Method

The sample consisted of twenty-one health educators (medical professors, nurses, occupational therapists, and health administrators) who attended a 1-week workshop in interpersonal communication and conflict sponsored by the Center for the Study of Medical Education, University of Illinois. The instructional techniques, developed by the National Training Laboratories, emphasized experience-based learning through competitive exercises, role playing, problem-solving tasks, and sensitivity training groups. An important goal of the workshop was to facilitate self-awareness. Past experience with similar groups at the Center suggested that health professionals, engaged with small peer groups, typically spend about 50 percent of their time giving opinions and expressing their wishes and only about 15 percent of their time responding to or seeking the ideas of others. In addition, self-ratings typically showed that they are unaware of these discrepancies: they rate themselves equally high on opinion giving and opinion seeking behaviors. Self-confrontation was used in this workshop to try to bring about more realistic self-perceptions.

Workshop participants were divided into two groups—A and B. Care was exercised to distribute MD's, nurses, health administrators, and females equally to make groups comparable in sex and professional background. Participants were video taped twice during the week in their respective T-groups. The major difference between groups A and B occurred in the timing and the mode of the feedback. Group A was exposed to focused video tape feedback twice, on Monday and Friday p.m.; both the group leader and participants stopped the tape repeatedly and discussed specific interactions portrayed in the tape (focused feedback). Group B reviewed both tapes on Friday p.m. but this review was *not* focused. The leader was present but neither he or the participants interrupted the tape; Group B members were permitted to draw their own conclusions (unfocused feedback).

The following diagram shows the design of the study:

	GROUP A (N=10)	GROUP B (N=11)
Monday a.m.	First video taping of T-group followed by self-rating	First video taping of T-group followed by self-rating
Monday p.m.	Focused review of first tape followed by self-rating	
Friday a.m.	Second video taping of T-group followed by self-rating	Second video taping of T-group followed by self-rating
Friday p.m.	Focused review of second video tape followed by self-rating	Unfocused review of first and second videotapes followed by self-rating*

*Both tapes were reviewed at the end of the week in Group B to also see whether sensitivity training alone, without self-confrontation, could bring about more realistic self-perceptions.

A short version of the Group Behavior Rating Scale (GBRS) was used to measure self-perceptions. The GBRS is based on Bales interaction analysis and has been used previously with similar groups (1). The short form consists of forty items and yields two scores: one for Reinforcer and one for Idea Seeker.[2] Ten items measure reinforcing behaviors and ten items idea seeking behaviors; the remaining twenty items are fill-ins to camouflage the purpose of the ratings. Spearman-Brown split-half reliabilities are quite high for both scores: .93 and .96 respectively, showing high internal consistency.

Subjects were instructed to recall their behavior with peer groups and rate themselves on a 10-point scale. A few sample items are shown in Figure 1.

Results

Does *focused* self-confrontation bring about more realistic self-perceptions? The answer to this question was obtained by comparing the Monday a.m. with the Monday p.m. and the Friday a.m. with the Friday p.m. self-ratings in Group A. Pertinent statistics are shown in Table 1.

Both the Monday and Friday focused self-confrontation produced a significant shift in self-ratings in the expected directions: after the tape review ratings were significantly *lower* on both scales. Group A participants realized that they engaged less frequently in reinforcing and idea seeking behaviors than originally supposed. Focused self-confrontation helped to adjust their perceptions to reality. (A measure of "reality" was available from observers' ratings of video tapes recorded during the Monday and Friday morning T-group sessions. These ratings showed that both reinforcing and idea seeking behaviors occurred no more than 15 to 20 percent of the time in both groups).

Although more reality oriented self-ratings were obtained after focused self-confrontation this change was temporary: Friday a.m. self-ratings were just as inflated as Monday a.m. ratings, in spite of informal feedback received from peers throughout the week, as part of sensitivity training. It seems that exposure to focused self-confrontation can produce an immediate reaction but its effects are short-lived. Lasting changes in self-perception can *not* be brought about by one-shot interventions. Repeated exposures to focused self-confrontation, similar to intermittent reinforcement schedules advocated by Skinner (8), seem necessary to strengthen and maintain a new, more realistic picture of the self.

Does *unfocused* self-confrontation bring about more realistic self-perception? Comparison of

Figure 1—Sample Items from 10-point Scale

REINFORCER	Almost never	Rarely	Occasionally	Usually	Almost always
Supports the ideas of others	1 2	3 4	5 6	7 8	9 10
Agrees	1 2	3 4	5 6	7 8	9 10
Refers to others opinions	1 2	3 4	5 6	7 8	9 10
IDEA SEEKER					
Seeks opinions	1 2	3 4	5 6	7 8	9 10
Asks for suggestions	1 2	3 4	5 6	7 8	9 10
Asks for possible ways of action	1 2	3 4	5 6	7 8	9 10

Table 1.—Comparison of Self-Ratings in Group A

	Monday a.m. before tape review		Monday p.m. after focused tape review			Friday a.m. before tape review		Friday p.m. after focused tape review		
Behaviors	Mean	SD	Mean	SD	t	Mean	SD	Mean	SD	t
Reinforcer	5.61	1.75	4.45	2.19	3.82*	5.91	1.75	4.74	1.91	2.10*
Idea Seeker	5.32	1.61	4.50	1.81	4.19*	5.47	1.76	3.91	1.90	2.14*

*Significant at the .05 level.

Table 2.—Comparison of Self-Ratings in Group B

	Friday a.m. before tape review		Friday p.m. after unfocused tape review		
Behaviors	Mean	SD	Mean	SD	t
Reinforcer	7.40	1.11	7.11	1.13	.31
Idea Seeker	7.24	.95	6.81	1.06	2.05

mean ratings showed that self-ratings in Group B remained the same after unfocused self-confrontation. This finding suggests that video tape review, without the leader's focusing remarks and subsequent discussion by participants, was not sufficiently provocative to bring about more realistic self-perceptions. Participants continued to overestimate the frequency of their responsiveness to others; they also overestimated their efforts to elicit suggestions and ideas from their peers.

Conclusions

The first important finding of this study is that self-confrontation had to be focused to help viewers attend selectively to cues that may be overlooked otherwise. If one goal of this workshop was to make health educators realize how rarely they engage in reinforcing and idea seeking behaviors, then confronting the group with as complex a stimuli as their taped interactions and letting them draw their own conclusions was *not* sufficient. They needed assistance to identify specific behaviors to be developed, just like Ivey and his associates (4) clearly identified *"attending behavior and reflection of feelings"* as two specific behaviors that beginning counselors needed to learn. Similar to other mechanical teaching aids self-confrontation appears to work best in addition to, and not instead of, teaching.

The second finding, frequently overlooked in studies based on only one experimental session, points up the necessity of repeated, focused self-confrontations to produce lasting changes. Skinner (8) was the first learning theorist to suggest that intermittent, periodic reinforcement is necessary to stamp in a new response. This study lends added substance to Skinner's findings. The sophisticated, adult participants of Group A changed their self-perceptions, which were developed over a long period of time, only temporarily under the impact of a single self-confrontation. Another study is planned with a similar sample during a 6-week workshop to learn more about the sequence and pacing of focused self-confrontations and their effects on self-ratings.

Although the major goal of this study was not to determine the impact of sensitivity training upon self-perception, data based on both groups clearly show that continued participation in T-groups did *not* lead individuals in this sample to realize how rarely they reinforce and seek the ideas of their peers. In both groups there were no significant differences between Monday a.m. and Friday a.m. mean self-ratings, although both groups were engaged in T-groups throughout the week. It appears that effects of informal verbal feedback on self-perception during sensitivity training merit further study.

FOOTNOTES

1. The authors are grateful to Mrs. Merry Omori and Miss Nancy Zega who reviewed and rated the video tapes.
2. Unpublished copies of both the long and short forms of the GBRS are available from the author upon request.

REFERENCES

1. Bales, R. F., *Interaction Process Analysis: A Method for the Study of Small groups*, Addison–Wesley, Cambridge, MA, 1950.
2. Boyd, H. S.; Sisney, V. V., "Immediate Self-Image Confrontation and Changes in Self-Concept," *Journal of Consulting Psychology*, 31:291–294, 1967.
3. Danet, B. N., "Self-Confrontation in Psychotherapy Reviewed: Videotape Playback as a Clinical Research Tool," *American Journal of Psychotherapy*, 22:245–257, 1968.
4. Ivey, A. E.,; Normington, D. J.; Miller, C. D.; Morrill, W. H.; Hass, R. F., "Microcounseling and Attending Behavior: An Approach to Practicum Counselor Training," *Journal of Counseling Psychology, Monograph Supplement*, 15:1–12 (No. 5, part 2), 1968.
5. Kagan, N.; Krathwohl, D. R.; Miller, R., Stimulated Recall in Therapy Using Video-Tape—A Case Study," *Journal of Counseling Psychology*, 10:237–243, 1963.
6. McGuire, F. L.; Steigel, T. T., "Psychotherapy by Television: A Research Paradigm," *Psychotherapy, Theory, Research, and Practice*, 3, 4, 1966.
7. Moore, F. J.; Chernell, E.; West, M. J., "Television as a Therapeutic Tool," *Archives of General Psychology*, 12:217–220, 1965.
8. Skinner, B. F., *Science and Human Behavior*, Macmillan, New York, 1953.
9. Stoller, F. H., "Focused Feedback with Video-Tape: Extending the Group's Function," in Gazda, G. M. (ed.), *Basic Innovations in Group Therapy and Counseling*, Charles C. Thomas, Springfield, IL, 1967.
10. Waltz, G. R., "Counselors Look at Themselves on Video-Tape," *Journal of Counseling Psychology*, 10: 232–236, 1963.

Research Briefs

THE SURPLUS OF PUBLIC SCHOOL TEACHERS: SOME EMPIRICAL ASPECTS[1]

Robert J. Thornton, Lehigh University

A RATHER unique phenomenon has developed in the labor market for public school teachers during the past 2 years—a substantial oversupply of qualified teachers. Estimates of the National Education Association indicate that more than 88,000 prospective entrants into the profession may have been unable to secure employment for the 1970–71 school year.

The advent of the surplus raises several empirical questions of both economic and educational interest. One such question concerns the effects on teachers' salary levels. Specifically, has the sudden turn of labor market conditions significantly affected the rate of change in teachers' salaries, as school boards find themselves more able to resist the economic demands of their staffs? A second and somewhat related question pertains to the response of school administrators in utilizing the surplus to improve staffing characteristics by effecting changes in (a) the proportion of teachers holding substandard certificates and (b) pupil–teacher ratios. The effects of the first 2 years of the teachers surplus with respect to both questions posed above were examined in this study.

Method

The measurement of the salary effects of the surplus was undertaken by developing an econometric model of the determination of teachers' salaries and by then fitting the model to cross-section data for the fifty states via multiple regression analysis. Although several alternative specifications of the model were tested, changes in average salaries (1969–70 to 1970–71) were hypothesized to be a function principally of: (a) the presence or absence of a general statewide surplus of teachers (measured dichotomously); (b) the average income level of the state; (c) selected educational revenue and expenditure variables, and (d) the extent of organized teacher activity in the state. For the purpose of estimating the effects of the surplus on differential changes in staffing characteristics, the relatively less sophisticated technique of tests of the significance of differences between sample means and proportions was used (t-tests). For the most part, data was secured from the National Education Association.

Results and Conclusions

Most of the regression coefficients for the variables in the salary-change equation failed to attain minimum levels of significance (10%) and the coefficient of correlation was disappointingly low ($R = .42$). However, there did appear to be a significant negative correlation between the presence of a teachers surplus and the percentage change in average salaries for the 1970–71 school year. Specifically, average salaries for teachers increased at a 2 percent lower rate in the surplus states ($t = 1.73$). As a percentage of the mean increase in salaries for all fifty states for this period (8%), this figure indicates that the advent of the surplus may have retarded the rate of increase in salaries by as much as 25 percent. On the other hand, there appeared to be no significant tendency for the surplus to be absorbed by declines in either the proportion of substandard teachers or pupil–teacher ratios for this period ($t = —.76$ and $.07$ respectively). The latter finding is, of course, not surprising in view of the increasing numbers of school districts recently faced with severe financial difficulties.

FOOTNOTE

[1] Based in part on a paper presented at the annual meetings of the Midwest Economic Association, April 1971.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 65, Number 2, October 1971)

Differences in Personality Traits Among Most Preferred and Least Preferred Students in Grades 10, 11, and 12

NORMAN J. BAUER

ABSTRACT

Subjects, 162 students in six classroom groups—two groups from each of grades 10, 11, and 12—were administered a comprehensive, 5-choice near-sociometric scale designed to identify most preferred and least preferred students. This variable, preference status, was then analyzed in relation to the personality characteristics measured by the California Test of Personality. Scores made by the most preferred students on each component of the personality test differed significantly (.01) from those made by the least preferred group. The preference status of students remains relatively constant through time.

IN ITS BROADEST sense, it is evident that education is growth toward socially desirable ends. Proper social development is of fundamental importance to all individuals. In fact, there is a growing conviction that guiding children and youth in the development of well-balanced, effective personalities constitutes the teacher's single most important function. Smith (3) points out that among the many aims for schools, two are outstanding: to help every child find personal security and to help the child develop the ability to function within groups.

These aims derive from the fact that every schoolroom is a social group and the relationships among the members form patterns similar to those in other groups. There are, as Northway (2) points out, the leaders, the cliques, the reciprocal companionships, and the isolated members. Because of the potential dangers isolated individuals hold for a democratic society, and because of the effects isolation has on personality development, it is mandatory that we concern ourselves with studies of the relative acceptability of students within their peer groups.

Almack (1) takes the position that association of all sorts is an ever-present fact of human existence, there is no choice so far as much of it is concerned. Everywhere we are dependent upon this fact for our comfort, happiness, and safety. Our days, he believes, are made up of face to face contacts with people, only the most intimate of which we do have any word in choosing.

This pervasiveness of patterns of interaction in social relationships has increased the importance of identifying factors related to social acceptability. Why is it that some individuals are highly accepted by their peers while others are neglected or even actively rejected by them? Are there distinct personality patterns that characterize the socially accepted and the socially rejected individuals? Answers to such questions may enable man to understand better the patterns of interaction which permeate social relations. Even more important, a study of the personality factors related to social acceptability and unacceptability should provide insights and suggestions for improving the status of those individuals who are socially unacceptable. The present study attempts to accomplish this by analyzing the relationships between preference status and the personality characteristics as measured by the California Test of Personality (CTP).

Population and Methodology

The study was conducted in a midwestern high school. Students (N = 162) in six classroom groups, two groups from each of grades 10, 11, and 12, were used in the study. The ages and intelligence scores of Ss were obtained from school records.

A scaled instrument was designed to measure the sociometric acceptability of the Ss in their respective peer groups. This was a comprehensive, 5-choice, near-sociometric scale, each category of which was assigned a numerical weight. The scores made by each student on this scale were computed. Those students in each group who had the ten highest scores on the scale were identified and together constituted the most preferred group of students. Those students in each group who had the ten lowest scores on the scale were identified and together constituted the least preferred

Table 1.—Self-Reliance

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	232.40	232.400	148.970
Between class levels	2	3.15	1.575	1.009
Interaction between class levels and groups	2	14.03	7.015	4.496
Within groups	114	178.00	1.560	
TOTAL	119	427.58		

group of students. In this way, sixty most preferred and sixty least preferred students were selected as the sample.

The CTP was used to measure the selected personality traits. This instrument was administered to the students, and scores were computed for each of the personality measures. These included measures of self-reliance, personal worth, personal freedom, belongingness, withdrawal, nervousness, social standards, social skills, anti-social tendencies, and family relations. An analysis of variance, multiple classification, was computed to determine the differences in the personality traits between the most preferred and least preferred groups of students.

One month after the initial sociometric scale was administered, the scale was readministered to each of the original 162 students. Correlation coefficients were computed between the scores made by Ss identified as most preferred or least

Table 2.—Personal Worth

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	261.07	261.07	109.23
Between class levels	2	11.22	5.61	2.34
Interaction between class levels and groups	2	.16	.08	.03
Within groups	114	272.86	2.39	
TOTAL	119	545.31		

Table 3.—Personal Freedom

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	193.54	193.52	80.64
Between class levels	2	8.60	4.30	1.79
Interaction between class levels and groups	2	65.46	32.73	13.63
Within groups	114	283.96	2.40	
TOTAL	119	551.56		

Table 4.—Belongingness

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	385.21	358.21	78.29
Between class levels	2	10.32	5.16	1.04
Interaction between class levels and groups	2	21.62	10.81	2.19
Within groups	114	561.28	4.92	
TOTAL	119	978.43		

Table 5.—Withdrawal Tendencies

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	163.33	163.330	49.640
Between class levels	2	1.51	.755	.229
Interaction between class levels and groups	2	174.73	87.365	26.554
Within groups	114	375.86	3.290	
TOTAL	119	715.43		

preferred on the first sociometric scale and the scores the same Ss received on the second sociometric scale.

Results

Tables 1 through 12 present an analysis of the differences in scores which were made by each preference group on each of the components of the standardized personality instrument used in this investigation.

Self-reliance, personal worth, personal freedom, a feeling of belongingness, withdrawal tendencies, and nervous symptoms constitute the Personal Adjustment category of the personality instrument. Social standards, social skills, anti-social tendencies, family relations, and commu-

Table 6.—Nervous Symptoms

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	106.41	106.41	21.49
Between class levels	2	43.12	21.56	4.35
Interaction between class levels and groups	2	12.02	6.01	1.21
Within groups	114	564.72	4.95	
TOTAL	119	726.27		

Table 7.—Social Standards

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	17.63	17.630	4.170
Between class levels	2	16.21	8.105	1.920
Interaction between class levels and groups	2	2.33	1.165	.276
Within groups	114	481.86	4.220	
TOTAL	119	518.03		

Table 8.—Social Skills

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	304.01	304.01	184.24
Between class levels	2	11.32	5.66	3.43
Interaction between class levels and groups	2	7.82	3.91	2.36
Within groups	114	219.12	1.65	
TOTAL	119	542.27		

nity relations represent the Social Adjustment category of the instrument.

Tables 13 and 14 summarize the analysis of variance computed to test for differences in mean total scores for each of these broad categories between the subgroups identified as most preferred and least peferred.

Significant differences between the most preferred and least preferred groups, at the .05 and the .01 levels of confidence, were found to exist on each component of the personality instrument. When the scores made by each class level were analyzed, a significant difference was found to exist on one component of the test, family relations. When the factor of interaction between grade level and preference group was considered, the analysis revealed a significant difference on

Table 9.—Antisocial Tendencies

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	218.69	218.690	89.260
Between class levels	2	.51	.255	.104
Interaction between class levels and groups	2	8.47	4.235	1.728
Within groups	114	279.50	2.450	
TOTAL	119	507.17		

Table 10.—Family Relations

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	371.01	371.01	508.23
Between class levels	2	18.82	9.41	12.89
Interaction between class levels and groups	2	51.52	25.76	35.28
Within groups	114	83.82	.73	
TOTAL	119	525.17		

Table 11.—School Relations

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	218.70	218.70	116.32
Between class levels	2	2.60	1.30	.69
Interaction between class levels and groups	2	5.60	2.80	1.48
Within groups	114	213.00	1.88	
TOTAL	119	439.90		

Table 12.—Community Relations

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	267.01	267.01	44.28
Between class levels	2	20.26	10.13	1.67
Interaction between class levels and groups	2	10.48	5.24	.86
Within groups	114	687.98	6.03	
TOTAL	119	985.72		

four components of the test, self-reliance, personal freedom, withdrawal tendencies, and family relations.

Discussion

These findings support the assertion that most preferred and least preferred adolescents differ significantly from one another on a number of presumably important personality traits. However, as with any finding of a relationship between measures of personality characteristics, the interpretation may not be clear. Further research needs to be conducted to determine whether individuals attain high status because they display certain personality traits or whether they develop these traits as a consequence of having attained high status.

Table 13.—Total Personal Adjustment Scores

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	8484.00	8484.00	421.380
Between class levels	2	38.50	19.295	.959
Interaction between class levels and groups	2	221.26	110.630	5.500
Within groups	114	2302.56	20.110	
TOTAL	119	11,046.32		

Table 14.—Total Social Adjustment Scores

Source of Variation	d.f.	Sum of Squares	Mean Squares	F
Between groups	1	8036.03	8036.030	405.040
Between class levels	2	214.71	107.355	5.411
Interaction between class levels and groups	2	123.23	61.615	3.104
Within groups	114	2262.66	19.840	
TOTAL	119	10,636.63		

REFERENCES

1. Almack, John C., "The Influence of Intelligence on the Selection of Associates," *School and Society,* 16:529–530, November 1922.
2. Northway, Mary L., "Outsiders: A Study of the Personality Patterns of Children Least Acceptable to Their Age Mates," *Sociometry,* 7:17, February 1944.
3. Smith, Gladys H., "Sociometric Study of Best Liked and Least Liked Children," *Elementary School Children,* 51:71–85, October 1950.

THE UNIVERSITY MIRROR

MANY students see the university as a mirror of the established order of things, an order which they would like to have changed. In that mirror, the university appears as an instrument of society created for the purpose of filling society's needs and hence conserving society's present methods, conditions, and dominating values. This view of the university is, of course, accurate.

But the critics of society's failures, even those whose frustrations over those failures have led them into violence and disruption, have not perceived that the university is also society's means for change and renewal. These functions do not show in the mirror, but they make possible the achievement of those social goals with which so many people are today strongly concerned. Hunger cannot be abolished without an abundant agriculture; poverty cannot be eliminated without successful business enterprise; ill health cannot be reduced without science; humane services cannot be performed without trained professionals; ideals for the improvement of society cannot be pursued effectively without freedom from the slavery of ignorance and from the restrictions of incompetence.—*President David Dodd Henry,* University of Illinois, at commencement ceremonies, Chicago, June 15, 1969.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 65, Number 2, October 1971)

Relationship of Study Habits and School Attitudes to Achievement in Mathematics and Reading[1]

FLORENCE P. SHEPPS[2] **and R. RONALD SHEPPS**
Kew Gardens Hills, New York

ABSTRACT

The Survey of Study Habits and Attitudes (SSHA) was used to examine the relationship of sixth grade study habits and school attitudes to achievement in mathematics and reading. Only the attitudes subtest appeared predictive for the group as a whole, using the reading criterion. More precise analysis of this data was accomplished through the use of a moderator variable procedure involving the grouping of pupils by sex. This analysis revealed that in the case of the boys only, the total SSHA score and the attitudes subtest predicted reading achievement. In the case of the girls, neither of these predicted the criterion, although the attitudes subtest did predict a different criterion, mathematics achievement. The implications of these findings are examined with respect to the early diagnosis of school difficulties as well as from the point of view of further development of the SSHA as a research instrument.

EDUCATORS vary in their opinion of the nature of good study habits and proper school attitudes. Evidence for this is demonstrated both in published works (2, 8) and in the wide differences that occur in the amount of homework assigned. Those who believe in home study and proper study habits are apt to assign heavy workloads. Other educators are more likely to favor proper motivation toward school as the key to effective learning. These individuals are apt to feel that study techniques are of decidedly secondary importance.

This controversy is probably due to the absence until recently of a satisfactory research instrument. (We are also facing a complex problem: It is likely that study practices and school attitudes are of different importance at various grade levels and with respect to different learning criteria.) The development of the Survey of Study Habits and Attitudes (SSHA) (2) gave researchers an instrument which appeared to be successful in attacking problems in this area (1, 5, 6, 7, 8).

The SSHA is a self-report type questionnaire which asks the student himself whether his place of study is kept neat, for example, or if study time is used efficiently and lessons learned in some planned order. The study questions also probe for other indications of the kind of habits which research has now established are associated with superior academic accomplishments (2). For example, information is sought on the test about such matters as whether teachers tend to be viewed critically, or the degree to which it is believed that grades are a result of teacher bias rather than accurately reflecting personal performance.

Research has established the validity of the SSHA as early as junior high school. Because of the importance of identification and elimination of faulty study habits and school attitudes at an even earlier stage, it was decided to try to use the SSHA on the sixth grade level. Would this carefully constructed, extensively validated instrument prove itself a valid predictor of academic accomplishments at this level?

The basic hypothesis tested in the present investigation was that it would. It was reasoned that sixth graders were not too dissimilar to seventh graders, so that reading difficulty with the test probably would not present a problem. Although junior high school students do tend to have a greater individual responsibility placed on them for the completing of assignments, it was thought that learning conditions in general were roughly comparable with those in the sixth grade, so that an instrument which predicted academic accomplishment as low as seventh grade could be expected to do so in the next lower grade.

Method and Procedure

Subjects for this study were sixteen boys and ten girls attending sixth grade at Parkway School, a private institution in Jamaica Estates, New York. Study habits and school attitudes were measured by the SSHA (2). Arithmetic achievement was measured by the Iowa Test of Basic Arithmetic Skills (3) and reading achievement by the Metropolitan Reading Achievement Test (4). Both achievement tests had been previously administered by the school and the students' scores recorded on school records.

Results and Discussion

Table 1 gives the correlation matrix for the two achievement criteria and all parts of the SSHA (habits, attitudes, total score). Rank order correlations were computed because they required no assumptions about equal interval measurement or an underlying normality of distribution.

Table 1.—Correlation of SSHA With Math and Reading Achievement (N = 26)

	Math	Reading	Habits	Attitudes	Total Score
Math		.25	.28	.29	.30
Reading			.25	.37*	.25
Study Habits				.76**	.88**
Attitudes					.54**

*Beyond .05 level of significance, one-tailed test, rank order coefficients
**Beyond .01 level

Table 1 indicates that although the several parts of the SSHA were substantially intercorrelated, only the attitudes subtest was a statistically satisfactory predictor of a criterion, reading achievement. More precise analysis of this data was accomplished through the use of sex group membership as a moderating variable. Tables 2 and 3 separately report the same correlation matrix for the sixteen boys and ten girls in the sample. Table 2 indicates that both school attitudes *and* total SSHA score predict the same reading achievement criterion among boys, while Table 3 shows that among girls only the attitudes measure is predictive, but for a different criterion, mathematics achievement.

The results suggest that the SSHA has some utility, at least as a research instrument, in the elementary school. The basic hypothesis of the study, that SSHA scores would be related to some criterion of academic performance of sixth graders, was supported in all analyses with respect to school attitudes. The study habits subtest was not predictive for either criterion, and the total score measure was predictive only in the case of male members of the sample.

The results point strongly toward the necessity of separate research examination in the elementary grades of male and female scores as well as separate analyses of the individual parts of the SSHA. Typically, these are not reported in the literature and the global criterion (grade point average) is favored over separate academic performance indices, although there are exceptions. In fact, the present test manual (2) itself fails to report separate male and female norms, although this had been done in earlier editions. The findings of the present study suggest that this is a mistake since failing to distinguish between scores of males and females was observed to obscure the relationship between SSHA total score and the reading criterion, in the case of the boys, and the attitude score and the mathematics criterion in the case of the girls. Evidently, the SSHA scores of elementary school boys, at least in samples similar to the present one, predict different performances than do the SSHA scores of girls.

Table 2.—Correlation of SSHA With Math and Reading Achievement (N = 16 boys)

	Math	Reading	Habits	Attitudes	Total Score
Math		.25	.15	.08	.18
Reading			.40	.60**	.50*
Study Habits				.78**	.87**
Attitudes					.95**

*Beyond .05 level of significance, one-tailed test, rank order coefficients
**Beyond .01 level

Provided the differences in the meaning of the SSHA scores of boys and girls are kept in mind,

Table 3.—Correlation of SSHA With Math and Reading Achievement (N = 10 girls)

	Math	Reading	Habits	Attitudes	Total Score
Math		.32	.38	.76**	.46
Reading			-.23	.02	-.33
Study Habits				.72*	.95**
Attitudes					.81**

*Beyond .05 level of significance, one-tailed test, rank order coefficients
**Beyond .01 level

research using several different criteria of academic performance should pave the way for the use of the SSHA in the elementary school for the early diagnosis of individual attitudinal problems hindering effective school performance. The instrument is also uniquely useful in that it offers the opportunity for item by item discussions in the course of counseling.

FOOTNOTES

1. This study represents a reanalysis and reformulation of material originally collected by the senior author in the course of writing her master's dissertation (8).
2. The senior author wishes to acknowledge the help of Dr. Jerome Radin in the collection and evaluation of the original data, and the kind assistance of Mr. Sidney Dorfman, Headmaster of Parkway School in Jamaica Estates, New York, in making available a subject population.

REFERENCES

1. Brooks, F. D.; Hestorn, J. C., "The Validity of Items in a Study Habits Inventory," *Journal of Educational Psychology*, 36:257–270, 1945.
2. Brown, W. F.; Holtzman, W. H., *Survey of Study Habits and Attitudes*, The Psychological Corporation, New York, 1956 and 1964.
3. Durost, W. N.; Bixler, H.; Hildreth, Gitl; Lund, K. W.; Wrightstone, J. W., *Metropolitan Reading Achievement Test* (Grade 6), Harcourt, Brace and World, New York, 1958–1961.
4. Lindquist, E. F.; Hieronymus, A. A., *Iowa Test of Basic Arithmetic Skills* (Grade 6), Houghton–Mifflin, New York, 1955–1956.
5. McGuire, C., *Talented Behavior in Junior High Schools*, The University of Texas, Austin, 1960.
6. Michael, W. B.; Reeder, D. W., "The Development and Validation of a Preliminary Form of a Study Habits Inventory," *Educational and Psychological Measurement*, 12:236–247, 1952.
7. Morris, F. L., "The Validity of the Brown–Holtzman Survey of Study Habits and Attitudes 1960 Experimental Revision for Grades 7–12," unpublished master's thesis, Wayne State University, Detroit, Michigan, 1961.
8. Shepps, F., "The Relationship of Study Habits and School Attitudes to Sixth Grade Achievement in Mathematics and Reading," unpublished master's thesis, Queens College, New York, 1969.

Forthcoming Articles

Emmer, Edmund T.; University of Texas at Austin
"TRANSFER OF INSTRUCTIONAL BEHAVIOR AND PERFORMANCE ACQUIRED IN SIMULATED TEACHING"

Goldman, Bert A. and Paul Vicinanza; University of North Carolina at Greensboro
"A SCORING SYSTEM FOR MEASURING COLLEGE INTEREST WITH THE KUDER OCCUPATIONAL INTEREST SURVEY FORM DD"

Grieve, Tarrance Don; University of Victoria, British Columbia and J. Kent Davis; University of Wyoming, Laramie
"THE RELATIONSHIP OF COGNITIVE STYLE AND METHOD OF INSTRUCTION TO PERFORMANCE IN NINTH GRADE GEOGRAPHY"

Hall, Lincoln H.; College of the Sequoias, Visalia, California
"PERSONALITY VARIABLES OF ACHIEVING AND NONACHIEVING MEXICAN-AMERICAN AND OTHER COMMUNITY COLLEGE FRESHMEN"

Haney, Jack N., William L. Albott, and Nicholas Paal; Ohio University, Athens
"SEX RELATED VARIABLES ASSOCIATED WITH CHOICE AND EVALUATION OF OPTIONS IN AN INTRODUCTORY PSYCHOLOGY COURSE"

Hardy, C. A.; North Texas State University, Denton
"ACHIEVEMENT AND LEVEL OF CRITICAL THINKING IN CHEM STUDY AND TRADITIONAL CHEMISTRY"

Hick, Thomas L.; State University of New York, New Paltz and Marilynn C. Santman; Frostig Center for Educational Therapy, Los Angeles
"TEST OF A STRATEGY TO INCREASE THE PREDICTABILITY OF FIRST GRADE READING SKILLS FROM LETTER NAMING ABILITIES IN KINDERGARTEN"

Null, Eldon J. and William H. Smead; Purdue University, Lafayette, Indiana
"RELATIONSHIPS BETWEEN THE POLITICAL ORIENTATION OF SUPERINTENDENTS AND THEIR LEADER BEHAVIOR AS PERCEIVED BY SUBORDINATES"

Osborn, Michael E.; University of Iowa, Iowa City
"THE IMPACT OF DIFFERING PARENTAL EDUCATIONAL LEVEL ON THE EDUCATIONAL ACHIEVEMENT, ATTITUDE, ASPIRATION, AND EXPECTATION OF THE CHILD"

Shaver, James P.; Utah State University, Logan and Dee Nuhn; Logan Public Schools, Logan
"THE EFFECTIVENESS OF TUTORING UNDER-ACHIEVERS IN READING AND WRITING"

Tilker, Harvey A.; San Diego State College, California and Kenneth M. Goldstein; Wakoff Research Center, Staten Island, New York
"ATTITUDES TOWARD A-B-C-D-F AND HONORS-PASS-FAIL GRADING SYSTEMS"

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 65, Number 2, October 1971)

The Use in Writing of Textbook Methods of Paragraph Development

RICHARD A. MEADE
University of Virginia

W. GEIGER ELLIS
University of Georgia

ABSTRACT

This study probed these questions: Which paragraph development methods suggested in textbooks are used in contemporary expository writing. Other than those suggested in textbooks, what methods of paragraph development are used in contemporary expository writing? What is the rate of occurrence of each of the employed methods of paragraph development? Composition textbooks published in the 1960's were examined to ascertain which known methods were to be sought. Samples of one hundred paragraphs each were taken from three sources. Each paragraph was analyzed to determine its dominant method of development. It was found that paragraphs were most frequently developed by examples, reasons, additional comment, or a variety of these methods.

THE PURPOSE of this study was to discover the extent to which methods of paragraph development are used in selected samples of contemporary expository writing. The answering of three questions led to the achievement of this purpose. First, which paragraph development methods suggested in textbooks are used in contemporary expository writing? Second, other than those suggested in textbooks, what methods of paragraph development are used in contemporary expository writing? Third, what is the rate of occurrence of each of the employed methods of paragraph development?

Procedures

Three major procedures were used in this investigation. The first was to determine the methods of paragraph development suggested in textbooks, the second to select paragraphs for analysis from current writings, and the third to analyze these paragraphs to determine (1) which of the textbook methods of paragraph development were used; (2) what paragraph development methods, other than those suggested in textbooks, were used; and (3) what the rate of occurrence for each of the employed methods of paragraph development was.

It is generally understood that textbooks strongly influence classroom content. An examination of English textbooks published in the 1960's revealed that one of the most common topics was methods of paragraph development. The following methods were discovered in the review: description, comparison, contrast, reasons, examples, definition, repetition, cause and effect, and chronology. Seven other methods were noted (arguments, details, facts, illustrations, restatement, order of space, and incident), but each appeared to be subsumed by one of the foregoing. The nine methods of development are defined as follows:

Description: Providing details that present a picture.

Comparison: Showing a similarity between at least two things.

Contrast: Showing a difference between at least two things.

Reasons: Answering the question *why* about the central idea. Cases in which the relationship is causal are excluded.

Examples: Citing a specific instance or case of something more general. Both being one of a number of specific cases and being a part of a single case of greater length qualify a statement for inclusion in this category.

Definition: Giving the meaning of a word or idea.

Repetition: Restating essentially the central idea.

Cause-Effect: Stating either end of a relationship which is presented as being causal.

Chronology: Presenting an occurrence as part of a larger event or as one of a series of steps in a process.

The paragraphs which were analyzed came from three sources: letters to a newspaper editor, professional journal articles, and magazine articles. These sources provided examples of writing from three different categories of writers: lay people, professional people in English education, and professional writers.

One hundred paragraphs for analysis were selected randomly from each of three sources: published letters to the editor of the *Richmond Times-Dispatch*, articles in *English Journal*, and articles in *Saturday Review*. Regular features and editorials were excluded so that the sample would not be dominated by a few writers. Since direct quotations may reflect the development methods used in the source, paragraphs comprised of such material were not included. The printed version of *Saturday Review* and *English Journal* paragraphs were used because the editors stated that they rarely change authors' indentations, but the original versions of paragraphs from the letters to the *Richmond Times-Dispatch* were examined so that editorial changes in content and paragraphing would not be reflected in the data.

Each of the three hundred paragraphs was studied carefully for categorization. Only paragraphs with a single central idea were categorized for predominant method of development (the eighteen paragraphs with two central ideas being so listed in Table 2). Paragraphs that showed no predominant method and paragraphs of only a single statement were also categorized separately (counted in Table 2 under the headings "A Variety of Methods" and "One-Statement Paragraphs" respectively).

Findings

The data show that over half the paragraphs were not developed by textbook methods: 62 percent of the letter paragraphs, 53 percent of the journal paragraphs, 53 percent of the magazine paragraphs, and 56 percent of the total three hundred paragraphs.

Development by examples was more frequent than any other textbook method in both the journals and the magazines, as was the case for all sources combined. Even in the letters, development by examples was a close second to development by reasons, which was the second most frequently used textbook method in all sources combined.

The only other textbook method of development which occurred more than 2 percent of the time was chronology. Paragraphs developed by description, contrast, definition, repetition, or cause-effect—the five other methods found—were relatively rare occurrences. None of the paragraphs was developed by comparison.

A chi-square analysis of the categories "textbook method used" and "textbook method not used" across the three sources of paragraphs revealed that the use or non-use of textbook methods of paragraph development was not associated with source—letters, journals, or magazines.

Source	Textbook Method Used	Textbook Method Not Used
Letters	38	62
Journals	47	53
Magazines	47	53

$x^2 = 2.18$; p = N.S.; df = 2

Three non-textbook methods of development were clearly identified: (1) opposition, (2) question, and (3) additional comment. A statement was judged as being development by "opposition" when it made an assertion in opposition to the paragraph's central idea. This method differs from "contrast" in that a contrasting statement asserts a difference between two things, whereas an opposing statement presents a conflicting position. A statement was judged as being development by "question" when it posed a question related to the paragraph's central idea. In the

Table 1.—Occurrence of Textbook Methods of Paragraph Development

	Method of Development										
Source of Paragraphs	Description	Comparison	Contrast	Reasons	Examples	Definition	Repetition	Cause-Effect	Chronology	No Textbook Method	Total Paragraphs
Letters	0	0	0	15	13	0	3	1	6	62	100
Journals	0	0	1	6	30	1	1	0	8	53	100
Magazines	1	0	4	7	27	1	0	2	5	53	100
Totals	1	0	5	28	70	2	4	3	19	168	300
Percent	.3	0	1.7	9.3	23.3	.7	1.3	1	6.3	56	99.9*

*.1 percent was lost by rounding each figure to the nearest tenth.

Table 2.—Occurrence of Paragraphs Not Developed by Textbook Methods

	Method of Development							
Source of Paragraphs	Two Central Ideas	Question	Opposition	Additional Comment	One-Statement Paragraph	A Variety of Methods	Total Paragraphs not Textbook Methods	Total Paragraphs
Letters	3	4	2	14	14	25	62	100
Journals	4	1	1	25	3	19	53	100
Magazines	11	0	4	19	1	18	53	100
Totals	18	5	7	58	18	62	168	300
							Total Percent	
Percent of occurrence among paragraphs not developed by textbook methods	10.7	3.0	4.2	34.5	10.7	36.9	100	—
Percent of occurrence among all paragraphs	6.0	1.7	2.3	19.3	6.0	20.7	—	56

method of "additional comment," statements were related to a paragraph's central idea in such a general way that they could not be characterized by any of the other methods. Rates of occurrence of these three non-textbook methods are given in Table 2.

As shown in Table 2, the "additional comment" and eclectic methods ("Variety of Methods") each account for about one paragraph out of every five, a rate much higher than that of any textbook method except exemplification. It would seem to be informative to break down further the sixty-two paragraphs developed eclectically. These paragraphs contain 184 supporting statements, which can be classified further as follows: reasons, 23; examples, 36; additional comment, 67.

Conclusions and Implications

It is clear that when the writers of the analyzed paragraphs developed those paragraphs, they most often did so by furnishing examples, giving reasons, making additional comment, or using a combination of methods, chiefly these same three. It is also clear, therefore, that textbooks neither emphasize these most popular methods nor include all methods which were observed.

What may composition teachers infer from these conclusions? If writings such as those examined are to serve as a guide for classroom instruction in composition, perhaps more and stronger attention to the nine traditional textbook methods will be advocated with the hope that young people now in school will in later life be certain to use them. Yet if the assumption is correct that textbook presence indicates curriculum content, then adult writers today did not lack for emphasis on methods of paragraph development during their own school days. Also, the chi-square test shows that even English teachers do not use textbook paragraph development methodology significantly more than do laymen. Therefore, it may be better to infer that, if formal instruction in paragraph development is in order, it should center on two traditional methods (examples and reasons), one new method (additional comment), and combinations of these.

It is conceivable that writers determine no method at all prior to writing but merely use a method which naturally follows from the main thesis of a paragraph. Research is yet needed to discover *how* writers make use of even these three methods of paragraph development. The research reported in this article reveals only *which* methods are most used.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 65, Number 2, October 1971)

Evaluation of an Innovative Educational Device–A Case Study[1]

DAVID P. YENS
University of Delaware and The Pennsylvania State University

ABSTRACT

A mechanical test-scoring device was evaluated in three ways: (1) experimentally to determine its contribution to classroom learning; (2) by questionnaire; and (3) by assessment of mechanical reliability. Televised learning materials were followed by three randomly assigned follow-up modes: (1) quizzes graded by machine, immediately returned, and discussed; (2) quizzes that were discussed, but with papers returned the next day; and (3) discussion only. Fourth and tenth grade classes were used. The value of such a study for the evaluation of devices for educational use was demonstrated. Machine malfunctions occurred in this target environment which had not appeared in extensive prior testing. The device groups showed no superiority over quiz only groups, possibly due to the reliability problem. Attitudes were favorable toward the machine and toward immediate return of the test results.

DURING THE past 10 years there has been a tremendous increase in the number and variety of mechanical and electronic devices aimed at the training and education market. A major impetus was given to the production of these devices by the Elementary and Secondary Education Act of 1965. Titles I and II of this act provided funds directly to schools and school districts for materials, aids, products, and programs to improve the education provided to the disadvantaged. The combination of federal support and aggressive salesmen led many school districts to order teaching machines, projectors, instructional television systems, and other devices regardless of the needs of the educationally retarded (2). Many of the devices sold had little proven educational value, were not sufficiently durable to withstand classroom use, or were not adequately supported with software.

The number of unevaluated educational devices on the market is still large and is continually growing. This places educators in the position of not being able to use a rational means of selecting a new device except on the basis of claims made by the manufacturer. Clearly, devices to be used for education or training should be evaluated in some way prior to marketing, and their validity for the purpose for which they are intended must be demonstrated. A method for performing this evaluation and its implementation with a new test scoring device is the focus of this paper.

Although there are a variety of test scoring devices currently on the market, they consist of either expensive electronic mark-sensing scorers or small self-scoring devices such as punch boards or those on which a mark between two lines changes color if marked correctly. At present, there seem to be no scoring devices that fit between these two extremes, although Pressey (7) described one almost 40 years ago. Little (5) tested the effect of the Pressey scoring machine on learning in a college educational psychology class and found that students who used the machine for immediate feedback from frequent quizzes performed better on a final examination than a control group that took the same quizzes but whose scores were returned after a day. However, these results did not lead to use of the machine at the time.

Because of the general similarity to the successful Pressey machine and because of the technological interests within the educational community, a recently developed mechanical scoring machine, called a "Scanalyzer," appeared to have potential as a valuable educational device. However, the effectiveness of a device at one educational level is not necessarily a guarantee of effectiveness at another level. In addition, the reliability of a device can be best evaluated in the environment in which it is to be used.

Method of Evaluation

A complete evaluation of an educational device should consider the following:

1. *Its actual contribution to some learning process;*
2. *Its acceptance by teachers, students, and school supervisory personnel;*
3. *Its reliability of operation or the consistency of its results.*

In order to meet these general requirements, the study was divided into two parts. The first part consisted of an experimental evaluation of learning incorporating the Scanalyzer machines compared with other methods of classroom practice. The second part comprised an assessment of attitudes by questionnaire and by personal discussion after the completion of the experimental phase. The third requirement was satisfied by insuring an adequate period of heavy use and by monitoring machine performance.

The experimental design that evolved to fit all requirements consisted of the following elements. The experimental evaluation would run for 9 weeks to afford sufficient machine usage. Grades 4, 7, and 10 were initially selected so that the suitability of the Scanalyzer for all grades could be tentatively estimated. Six classes in each grade were considered the minimum acceptable number for analysis; practically, it was expected that this number would be the maximum attainable. Three modes of activity were selected to follow standardized presentations in each grade: a quiz graded by the Scanalyzer, a normal quiz, and discussion only.

The need for a standardized subject matter was filled by using programs broadcast by WPSX–TV, the educational television station of The Pennsylvania State University, which were used by a number of classrooms in the central Pennsylvania area. This permitted the simultaneous viewing of the same material at many locations and simplified the logistical problems of providing materials to the participating teachers.

Television programming covered a wide range of grades. Once the grades to be used for the experiment were selected, programs at those levels were reviewed to find the ones most suitable. Series of programs were selected based upon their appropriateness to the study, their availability for previewing, the novelty of their contents to the students, and the extent of their use.

Three series of programs were selected: "AIBS Biology" for the tenth grade, "Profiles in Courage" for the seventh grade, and "In the News" for the fourth grade.

Classes were selected on the basis of their established viewing habits, the travel required to reach them, and the willingness of teachers and school administrators to participate. Classes that watched the selected program on a regular basis were sought to minimize the impact of the experimental treatment on the class and minimize the corresponding "Hawthorne effect" (8). It was considered desirable to make the students think that the testing was associated with an improvement in the use of the televised programs rather than being part of an experiment since the feeling of being part of an experiment might engender abnormal behavior on the part of the students.

One high school requested that all its biology and social studies classes be used, which provided a total of twelve classes for biology and fourteen for social studies. Due to scheduling difficulties, this later resulted in "Profiles in Courage" being used in tenth grade American History courses rather than in junior high classes as desired. Unfortunately, this difficulty arose after it was too late to locate other classes that would be more suitable. However, the increase in the number of classes was expected to provide a better trial of the machines and would permit the statistical analysis to be more powerful.

Subjects. The initial composition of the sample is indicated in Table 1. Teacher B dropped out later at his request. The class sizes ranged from ten to thirty-seven, but most were between twenty-five and thirty-five. They comprised an ability span from below average to markedly above average, and included most socioeconomic groups. Rural areas, small towns, and medium-sized towns (population 10,000) were represented. All the schools lie in the Appalachia region, and probably could not be strictly considered typical of rural or small-town America.

Most of the teachers had had a great deal of teaching experience and seemed to be highly competent.

Equipment. The Scanalyzer is a mechanical machine that is hand operated by rotating the crank on the right side. It uses special partially-prepunched data processing cards ("Port-a-Punch" cards) that can be perforated with any sharp instrument, such as a pencil.

A master card, punched with the proper answers, is inserted into a slot on the machine and is locked in place. A student punches his answers in another card, inserts it in another slot, and rotates the crank to score his card. The machine simultaneously compares the punches in each column by an arrangement of sensing pins. If all punches in the column match, it punches a hole at the top of that column. The number of questions answered correctly is the total number of holes punched in the top of the card. In addition, for each correct answer a corresponding dial on the face of the machine is rotated one place. These dials are additive so that after scoring a class the teacher would know how many students got each question correct. Knobs on the right-hand side of the machine allow resetting the accumulator dials.

The machine weighs 15 pounds and is completely portable. It can score ten 5-choice questions at a time, but there is a provision for discriminating between different sets of cards which would allow the scoring of cards used sequentially for a longer test. Ten prototype Scanalyzers were furnished for the study.

Materials. Quizzes for every televised program were developed by the project staff. In order to minimize the ceiling effect and insure class discussion, the tests were made reasonably difficult (most students judged them average to hard). In order to use the multi-answer capability of the Scanalyzer and increase test difficulty, many questions were constructed that required more than one correct answer. All questions could be an-

Table 1.—Initial Distribution of Teachers and Classes

TV Program	School	Teacher	No. of Classes
AIBS Biology (tenth grade)	H.S. #1	A	6
	H.S. #1	B	5
	H.S. #2	C	1
Profiles in Courage (tenth grade)	H.S. #1	D	6
	H.S. #1	E	5
	H.S. #2	F	3
In the News (fourth grade)	County School System	G	2
	(each teacher at a different school)	H	1
	"	I	1
	"	J	1
	"	K	1

[1] H.S. = High School

swered after having watched the program. The questions were based on previews of the program material as well as on the collateral material that accompanied both "AIBS Biology" and "Profiles in Courage." Because of a short time between production and broadcast of "In the News," the project staff generally sat in on the production itself.

"In the News" was produced by WPSX-TV. Each 15-minute program was concerned with a different aspect of current news and events. Topics included the Winter Olympics, the President's State of the Union message, hippies, and the plight of our cities.

"Profiles in Courage" was a dramatic series based on the book by John F. Kennedy. It was originally produced and shown by NBC. Each program was an hour in length, but could be divided into half-hour segments. These programs dealt with critical events requiring courage by some of America's great, but sometimes little-known, historical figures.

"AIBS Biology," a series of half-hour programs produced by the American Institute of Biological Science in cooperation with a text-film firm, was concerned with special topics in biology and included lectures, demonstrations, and many visuals.

Experimental Design. The experiment was designed to run for 9 weeks. An additional 2-week period was required at the end for final posttests and for administering the questionnaires.

Three modes of post-program activity (description follows) were sequentially used in each class, with each mode requiring 3 weeks. The sequence of modes for each class was randomly determined.

For each televised series, the following modes were used:

1. Scanalyzer mode. *Classes were given quizzes after each televised presentation, the quizzes were immediately graded with the Scanalyzer, and a following discussion of the quiz was directed toward those questions that tended to be missed by the class. Students usually scored their own cards. The score cards were retained by the students during class, but all cards were turned in at the end of the class for later analysis.*
2. Quiz mode. *Classes were given the same quizzes as the Scanalyzer groups after each televised presentation, but these were turned in, graded by the teacher, and returned the next day. Normal discussion of the quizzes followed their completion.*
3. Control mode. *The televised material was discussed after the presentation; no quizzes were given. The teachers were given copies of the quiz that was given in the other classes as a guide to the discussion.*

Experimental materials were hand-delivered each week to teachers and schools needing them. This permitted weekly evaluation of machine performance and maintained close contact with the participants or schools.

The extent of prior knowledge about program content was assessed for "AIBS Biology" and "Profiles in Courage" by a pretest before each 3-week period. The questions were drawn from the program content and were afterwards modified and included in the weekly quizzes. No pretest was possible for "In the News" due to the short time between planning, production, and broadcasting. A first attempt at a pretest for this program was disregarded due to a programming change.

Each 3-week period was followed by a posttest which was composed of the same questions as the pretest, but in a scrambled order. The posttests for "In the News" consisted of revisions or combinations of questions from the previous weekly tests. Because of the nature of the programs, most of the "In the News" material could be considered new to the students.

Test reliabilities were determined with the Kuder-Richardson Formula 20 (3). It would be expected that a pretest consisting of novel responses would yield a low reliability because of the high error component due to guessing. A sample of the pretests administered indicated that this was the case: pretest reliabilities for the biology tests ranged from .12 to .33 and for "Profiles in Courage" from .10 to .36. The obtained reliabilities for the posttests were somewhat low: .38 to .45 for the biology test; .57 to .67 for "Profiles in Courage"; and .52 to .71 for "In the News." Because the tests were short, were like "teacher-made" tests, and were designed to be difficult, the reliability coefficients for the latter two program tests are considered acceptable. Those for the biology tests are lower than desirable and indicate a more serious lack of internal consistency due to a moderately high error component. Results should be considered in this light.

All master cards, blank cards, tests, and other required materials were provided to the teachers. For programs with content that might be unfamiliar to the teacher, a summary of the program material was prepared and distributed to teachers.

The purpose and conduct of the experiment was discussed thoroughly with each teacher prior to participation, and all pertinent information was provided in a handout which contained detailed information about the experiment, the operation of the Scanalyzer, and directions concerning scoring and record keeping. Special forms were prepared to aid the teachers in recording the scores.

In addition to the formal part of the experiment, teachers were encouraged to use the Scanalyzers on their own or to allow their use by others. They were supplied with additional cards for this reason. Some teachers took advantage of this.

The *a priori* expectations of the experiment were:

1. *Scanalyzer groups would perform better than quiz groups.*
2. *Scanalyzer and quiz groups would perform better than the control groups. (This was a check on the quizzes themselves, since there is abundant evidence that quizzes do improve recall (cf 6 and 9).)*
3. *There would be no difference in performance among the three subject matter groups.*
4. *There would be interaction between modes and posttest performance and/or gains.*
5. *There would be no interaction between modes and subject matter programs.*

Questionnaire. Following the last posttest of the experimental phase, a questionnaire was administered to all students and teachers to survey their attitudes toward the Scanalyzer, the experiment, and the tests. Students completed the forms anonymously. Different questionnaires were designed for the teachers, the elementary students, and the high school students. A Likert scale was used for responses by teachers and high school students. Additional questions on the teacher form required short answers. Written comments were solicited for all questions. Surveys for the fourth grade students consisted of yes-no type questions with added space for comments.

Operational Difficulties Encountered. Because of scheduling, class duration, and program duration, classes differed in their use of televised and experimental materials. The elementary classes watched the programs and were able to complete the subsequent activities as planned. However, in the high schools class periods were 40 or 45 minutes long, and the "AIBS Biology" and "Profiles in Courage" programs were 30 minutes long, leaving only 10 or 15 minutes for the quiz, grading, and discussion. This problem was handled in two ways; for most biology programs, the test was given immediately after the program and was graded if appropriate. The discussion of the quiz was then held the next day. Because classes in the quiz mode were to have their scores returned to them the next day, biology students in both Scanalyzer and quiz modes generally knew their results at the time of the discussion and any advantage conferred by the Scanalyzer might have been minimized.

This problem was minimized for "Profiles in Courage" by showing the programs jointly to all participating classes during an activity period first thing in the morning. The hour-long program was divided into two portions and was shown on successive mornings. The appropriate activity (quizzing, grading, discussion) was then carried out during the normal class period on the day of the last viewing. However, different times elapsed between the presentation of the stimulus material and the quizzes, artificially increasing the variability between classes and possibly reducing the difference between the modes of activity.

Occasional operational difficulties were encountered with most of the Scanalyzers. Malfunctions were repaired weekly, but due to the heavy use of the machines interim measures had to be used by some classes that were assigned to the Scanalyzer mode.

Results

Statistical Analysis. The overall statistical design contained four factors: (1) the three television programs; (2) the three modes (Scanalyzer, quiz, and control); (3) the three trials (the 3-week periods); and (4) the pairs of pretests and posttests. Due to the statistical difficulties inherent in the analysis of gain scores (1, 11), pretest and posttest scores were treated as levels of one factor in analyses of variance.

The design was complicated by the nesting of teachers under programs and of classes under teachers, and by the fact that modes and trials are not completely balanced. Thus, it was necessary to split the data in different ways and use a variety of statistical analyses in order to test the significance of each factor.

At the outset, it is necessary to observe that the results of the analyses must be interpreted with extreme caution due to the variations in the design and operational difficulties described above. Because of the lack of true randomization and because differences between the different programs were tested with a Lindquist VI design (4), assumptions implied by the analysis of variance tests were not met. Thus, these results should be considered only as indications of trends in the data rather than as results of meaningful tests of significance.

Table 2.—Biology Classes: Mean Pretest and Posttest Results and Gains by Mode and Trial. N is the Number of Contributing Classes

Mode	Trial 1				Trial 2				Trial 3				Mode Means		
	N	Pre-test	Post-test	Gain	N	Pre-test	Post-test	Gain	N	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain
Scanalyzer	2	3.09	6.73	3.64	1	1.73	2.17	0.44	4	2.64	4.76	2.12	2.64	4.95	2.31
Quiz	4	2.99	7.49	4.50	3	1.89	3.44	1.55	0	-	-	-	2.52	5.75	3.23
Control	1	3.22	5.38	2.16	3	1.76	3.34	1.58	3	2.53	4.07	1.54	2.30	3.94	1.64
Trial Means		3.05	6.97	3.92		1.81	3.21	1.40		2.59	4.46	1.47	2.49	4.88	2.39

Because intact classrooms were used, the basic unit of analysis was the class mean. Thus, the following means represent the average of class means. All tests consisted of ten questions. Pretest and posttest means and gains of the high school classes are given in Tables 2 and 3 by trial and by mode. Table 2 contains these data for "AIBS Biology" classes. The sample consisted of two teachers and seven classes. Table 3 contains data for "Profiles in Courage" classes. The sample consisted of one teacher and three classes. In both tables, each trial by mode cell may represent a different number of classes, indicated by N. Posttest scores for "In the News" are in Table 4. In this table five teachers and six classes are represented. The number of classes contributing to each cell is specified by N.

Two teachers using "Profiles in Courage" were dropped from the analysis because the procedures they used did not appear to correspond with the planned design. One biology teacher dropped out at his own request. Due to the *a priori* randomized assignment of classes to modes by trial cells, this loss resulted in some unfilled cells as well as unbalanced cells. The mode means and the trial means are based on the same number of classes, however, and provide the best comparison between the different elements.

The differences among the three mode means and among the three trial means are readily apparent for all programs. All high school classes but one (a biology quiz class) gained from pretest to posttest. Using a two-way analysis of variance, the differences between the test means were

Table 3.—"Profiles in Courage" Classes: Mean Pretest and Posttest Results and Gains by Mode and Trial. N is the Number of Contributing Classes

Mode	Trial 1				Trial 2				Trial 3				Mode Means		
	N	Pre-test	Post-test	Gain	N	Pre-test	Post-test	Gain	N	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain
Scanalyzer	0	-	-	-	2	1.62	7.67	6.05	1	3.29	5.17	1.88	2.17	6.83	4.66
Quiz	2	1.39	5.06	3.67	1	1.92	6.76	4.84	0	-	-	-	1.57	5.62	4.05
Control	1	1.21	4.66	3.45	0	-	-	-	2	4.28	5.56	1.28	3.25	5.25	2.00
Trial Means		1.33	4.92	3.59		1.72	7.36	5.64		3.95	5.42	1.47	2.33	5.90	3.57

Table 4.—Mean Performance of Fourth Grade Classes on "In the News" Posttests, by Mode and Trial. N is the Number of Contributing Classes

Mode	Trial 1		Trial 2		Trial 3		Mode Means
	N	Mean	N	Mean	N	Mean	
Scanalyzer	2	4.98	1	4.16	3	4.72	4.83
Quiz	2	4.33	2	6.83	2	4.39	5.18
Control	2	2.67	3	3.96	1	6.70	3.99
Trial Means		3.99		4.95		5.06	4.67

significant (p<.01) for both programs. No significant differences were found between the pretests (one-way analysis of variance). All fourth grade classes performed better than a fifth grade class that had no contact with the televised programs, indicating that learning took place.

In general, the data show that the quiz groups were superior to the other two on the posttests and the gains, and the control groups were almost always lowest. However, there were no significant differences among mode means within programs. The one program in which the Scanalyzer groups did better than the other groups was "Profiles in Courage," but this group consisted of only one teacher and three classes. Thus, the "Profiles" groups cannot be considered as meaningful as the other programs which had at least twice as many teachers and classes. Furthermore, two of the "Profiles" Scanalyzer groups occurred in the second trial in which all posttest scores were high and pretest scores were low, and two of the quizzes were graded by hand due to machine malfunctions. If this was an indication of an easy test, then the final comparison was artificially inflated due to the heavy weight of this random occurrence or due to the modified experimental conditions. More classes would have been needed to allow a more definite interpretation.

The possibility of improved posttest performance being the result of external learning during the pretest-posttest interval was evaluated by giving the posttest to an additional, unrelated control group that had not been pretested. The posttest scores of all "Profiles" classes were superior to those of a nonparticipating senior English class, although the English class mean was somewhat higher than the pretest means of the social studies classes.

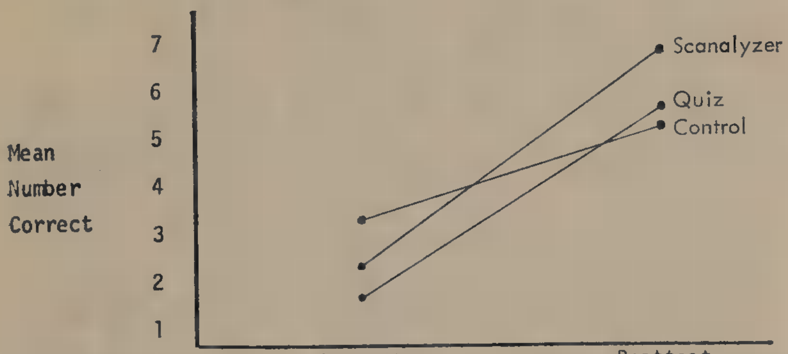


Figure 1—Test by Mode Interaction for "Profiles in Courage" Classes

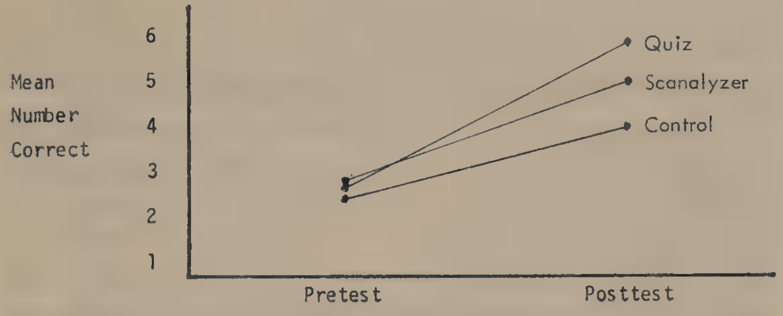


Figure 2—Test by Mode Interaction for Biology Classes

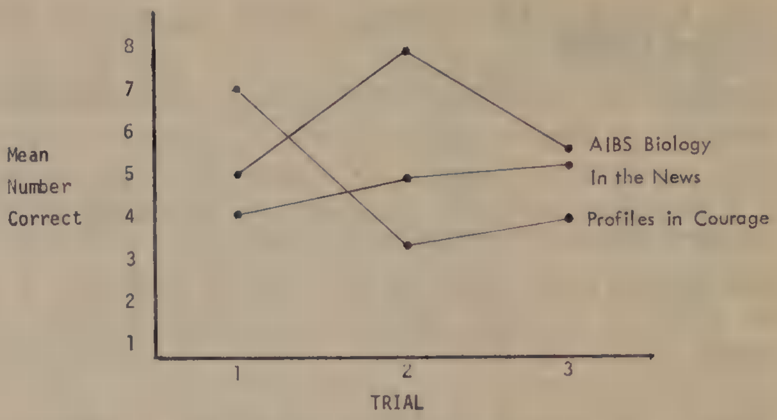


Figure 3—Changes in Posttest Means by Trial for each Program

An illustration of the differences described previously is found in Figures 1 and 2, graphs of the gains for each mode group for the "Profiles" and biology classes, i. e., the modes by tests interaction. This interaction was not significant (p>.05) for either program, contrary to expectations, indicating a similar pattern of gains for each mode.

The pattern of differences among trial means differs for the three programs. The differences among posttest means for each trial are shown graphically for each program in Figure 3. The obtained gain scores follow a similar pattern. The progressive increase for "In the News" classes indicates that either the students became progressively better at learning from television or that the difficulty of the tests decreased. (Teacher comments supported the former possibility.) The absence of pattern for the other programs indicates differences in the difficulty of the materials or tests, or perhaps in class attitudes toward the materials or the experiment.

An analysis of the "Profiles in Courage" and biology programs using a Lindquist VI design (4) was made with the understanding that unlike elements (programs) were involved. However, some information concerning this factor was desired. No significant differences were found between programs (F = 4.579 with 1 and 8 degrees of freedom, p>.05). The interaction between programs and modes was not statistically significant, but the programs by tests interaction was significant (F = 13.517, p<.01), reflecting the greater gains made by the "Profiles in Courage" classes. The modes by tests interaction for this analysis was not statistically significant (F = 3.044 with 2 and 16 degrees of freedom, p>.05).

A review of the expectations set forth earlier yields the following observations, based on the results of the experiment:

1. *It was not found that Scanalyzer groups perform better than quiz groups.*
2. *Scanalyzer and quiz groups did perform better than control groups, but the difference did not reach statistical significance. All groups did learn, as indicated*

by a significant difference between pretests and posttests.
3. *No difference between two program groups was found.*
4. *No interaction between modes and tests was found.*
5. *No interaction between modes and programs was indicated.*

It is necessary to note again that these observations are highly tentative and must be interpreted with caution due to the variations in design, low test reliabilities, and the malfunctions of the Scanalyzer.

Questionnaire. For each program a tabulation of the students' responses to each question was made and a summary of all the comments and answers was prepared. A summary was provided over all teachers.

Extensive information concerning opinions about the machine and experiment as well as attitudes toward the use of immediate feedback and the role of a device such as the Scanalyzer in the classroom was obtained from the questionnaires. In general, students and teachers complained about the lack of reliability of the machine, but liked the idea and the concept of immediate return of the tests. A machine like the Scanalyzer would be useful if it worked properly, and it would be a real aid to learning, according to the responses. Almost all the students said they liked to have their quizzes quickly marked and returned.

The teachers and students in the elementary schools were more enthusiastic about the machine than their counterparts in high schools, but this was probably because they suffered fewer malfunctions and had more time in which to accomplish the quizzing and discussion called for by the experiment. The fourth grade teachers pointed to a possible supplemental benefit of the quizzes: administration after each program increased the students' attention to the program and taught them to listen more carefully. Four teachers used the Scanalyzer on their own for grading class quizzes. They generally were pleased with their results.

Extra Experimental Use. In addition to personal use by teachers participating in the experiment, the Scanalyzer was also successfully used by another teacher with two high school special education classes, one in history and one in English, for 2 days each. Students were assigned to these classes on the basis of exceptionally low IQ scores. Testing was normally quite limited because the students could not remember the questions when their results were returned after a day. For a test to be effective, this type of student needed his results right away. Both the teachers and the students were pleased with the Scanalyzer and felt that its use was successful. The teachers thought the machine could be used for all the testing they do; its use might encourage more frequent testing.

Mechanical Malfunctions. All but one Scanalyzer developed some sort of malfunction during the study. Most appeared to be the result of internal changes in alignment caused by the rugged use given the machines or due to improper cleaning or use of the machines. However, the cause of some difficulties could not be determined. Most disorders were repaired by the project staff on a weekly basis.

The major result of the lack of reliability of the Scanalyzers was that students did not trust the results of tests scored by the machines, so machine scores were checked by hand scoring afterward. A detailed description of all difficulties encountered and corrective measures used was included in the final report (12) to the sponsor.

Administrator Interviews. Interviews with high school administration personnel revealed some hostility toward the study. Those interviewed said that there had been high enthusiasm for the Scanalyzer at first, but mechanical difficulties and the excessive amount of class time required had fostered increasingly negative attitudes toward the experiment. However, they maintained that the idea of a mechanical scoring device was good and they made recommendations for improvement. They stressed the need for very high reliability in a machine of this type, and discussed the administration's attitude toward the balance of costs and benefits.

Discussion

The value of fully evaluating a new educational device in a natural environment has been demonstrated by the results of this study. The importance of a study such as this one is shown by a number of factors. First, a prototype Scanalyzer identical to those used in the schools was cycled more than 5,000 times by a secretary in an office. No malfunctions occurred. It was estimated that the machines used most in the schools were cycled between 3,000 and 4,000 times, yet a variety of malfunctions developed. Apparently, some differences exist between classroom and office use of the machine that indicate the need for testing in the classroom.

Second, in spite of the poor reliability of the Scanalyzers, the questionnaire results indicated a desire by teachers and students for a machine of this type used in the ways specified. This would be valuable information for a potential manufacturer. In the absence of such information it would be difficult to make a decision whether to improve the machine or discontinue development. Without attitude information, a producer might feel that the initial investment would justify further development. If there had been a latent unfavorable response and machines were sold to schools by high-pressure salesmen, in all probability they would not be used by teachers and the investment by the school systems would be wasted. Thus, the assessment of attitudes, with an appropriate response by a producer, should tend to insure that only worthwhile devices that would be utilized by teachers and students would be offered to school systems.

Finally, although beset by a variety of difficulties, the feasibility of an experimental evaluation of the contribution to learning by an educational device has been demonstrated. It was possible to find teachers and administrators who were willing to try a new device in a certain way and who would devote class time to a program, testing, and discussion, and personal time to record keeping. The educational television programs proved to be an excellent way of providing standard learning materials to classes at different locations at the same time. Motion pictures could be used for the same purpose, but if they were not shown at the same time the time difference would be an added uncontrolled variable.

A number of factors confounded the experimental results of this experiment which indicate where improvements could be made in future studies of this type. Clearly, the lack of reliability of the prototype Scanalyzers themselves was the major disruptive element. For new devices, this is, perhaps, the factor over which the least control can be exerted. As has been seen, a device can be tested and proven in an ideal environment but fail to perform properly in the environment for which it is ultimately intended. A solution would be to place devices in schools for a trial test period of about 2 months with some incentive for teachers to use them heavily. All malfunctions encountered should be corrected by design changes and retested prior to initiating the experiment. An alternative would be to attempt to simulate the real environment and give prototype machines at least twice as much use as would be normally expected in 2 months. Again, design modifications would be preferred to simply fixing the machines prior to experimental use. The alternative to these solutions is to use untested prototypes and hope that nothing goes wrong. The proposed solutions would add some months to the evaluation process, but would markedly improve the prospects for a successful experimental evaluation.

Recommendations

The procedures described in this paper, with the suggested modifications, are recommended for use by manufacturers of new educational devices prior to marketing or by educational systems or institutions prior to purchasing new, untried educational devices. Only in this way is it possible to assess the true value of innovative educational devices and insure the best use of scarce funds for education.

Although the evaluative process was tried out with an educational device, it might be possible to use analogous designs for the formative or summative assessment of new educational systems and curricula. Most curricular improvements are tested by simply implementing them in one or more schools and determining if specified objectives are attained, or by conducting a "horse race" between the new curriculum and a traditional curriculum. A recent review by Welch (10) indicates that counterbalancing of treatments and the inclusion of other competing curricula have not generally been done. Systematic evaluation of teacher and student attitudes has not usually been performed. Use of more comprehensive designs, such as the one described in this paper, might lead to improvements in the assessment of the many new educational products, systems, and curricula that are being introduced at the present time, and may lead to a more efficient, more productive, and better liked system of education than the one we have now.

FOOTNOTE

1. The major support for this project was provided by a grant from IRC, Inc. The author wishes to thank John Wynne, Dr. H. Dennis Sherk, and Dr. Harold E. Mitzel for their assistance with this study.

REFERENCES

1. Cronbach, L. J.; Snow, R. E., *Individual Differences in Learning Ability as a Function of Institutional Variables, Final Report*, California School of Education, Stanford University, Palo Alto, March 1969.
2. Drew, E. B., "Education's Billion-Dollar Baby," *Atlantic Monthly*, 218:37–43, January 1966.
3. Kuder, G. F.; Richardson, M. W., "The Theory of the Estimation of Test Reliability," *Psychometrica*, 2:151–160, 1937.
4. Lindquist, E. F., *Design and Analysis of Experiments in Psychology and Education*, Houghton Mifflin Co., Boston, 1953.
5. Little, J. K., "Results of Use of Machines for Testing and for Drill upon Learning in Educational Psychology," in Lumsdaine, A. A.; Glaser, R. (eds.), *Teaching Machines and Programmed Learning*, National Educational Association, Washington, D. C., 1960, pp. 59–65.
6. McKeachie, W. J., "Research on Teaching at the College and University Level," in Gage, N. L. (ed.), *Handbook of Research on Teaching*, Rand-McNally and Co., Chicago, 1963, pp. 1118–1172.
7. Pressey, S. L., "A Third and Fourth Contribution Toward the Coming 'Industrial Revolution' in Education," in Lumsdaine, A. A.; Glaser, R. (eds.), *Teaching Machines and Programmed Learning*, National Educational Association, Washington, D. C., 1960, pp. 47–51.
8. Roethlisberger, F. J.; Dickson, W. F., *Management and the Worker*, Harvard University Press, Cambridge, MA, 1939.
9. Rothkopf, E. Z.; Bisbicos, E. E., "Selective Facilitative Effects of Interspersed Questions on Learning from Written Materials," *Journal of Educational Psychology*, 63:56–61, 1967.
10. Welch, W. W., "Curriculum Evaluation," *Review of Educational Research*, 39:429–443, 1969.
11. Wodtke, K. H., "On the Assessment of Retention Effects in Educational Experiments," *Journal of Experimental Education*, 36:(no. 4)28–36, 1967.
12. Yens, D. P.; Wynne, J. P.; Sherk, H. D.; Mitzel, H. E.; *Final Report: An Evaluation of the Scanalyzer in Elementary and Secondary School Classrooms*, College of Education, The Pennsylvania State University, University Park, September 1968.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 65, Number 2, October 1971)

Open-Closed Mindedness and an Inquiry-Oriented Social Studies Methods Course[1]

JAMES P. SHAVER
Utah State University

HYRUM E. RICHARDS
Casper—Midwest Schools, Casper, Wyoming

ABSTRACT

A social studies methods course focused on preparing teachers to teach students skills of critical thinking was studied. The course concentrated on the examination of assumptions underlying social studies curricular decisions and on critical thinking skills. Discussion, rather than lecture, was the primary method of instruction. The Dogmatism, F-, and Rigidity Scales were administered as measures of open-closed mindedness. Correlations were run to determine if scores on the three scales were related to ratings of the instructor by the students, using the University of Washington Survey of Student Opinion of Teaching, or to achievement in the course. Scores on the personality measures showed no consistent relationship with ratings of the instructor. There also were no significant relationships between the personality measures and the measures of learning. With a larger sample, many of the correlations would have been statistically significant, even though the relationships were educationally trivial.

SOCIAL STUDIES educators have long claimed that critical or reflective thinking is a central goal of social studies instruction in the public schools. Nevertheless, training programs for social studies teachers have not typically included courses that would give the prospective teacher a conceptual framework from which to teach his students to think in a more systematic manner (4, 38). Logic courses are, for example, usually not required in teacher preparation programs, and there are few other courses at the university that focus on thought process rather than content.

The social studies curriculum and methods course universally required of prospective social studies teachers provides one place where instruction might focus on critical thinking skills to be taught in social studies. In order to use the course for that purpose, however, the instructor would have to abandon some traditions of teacher education—for example, that a considerable portion of the methods course should be spent in training prospective teachers to plan lessons and units. Moreover, if the prospective teacher is to properly integrate new analytic concepts into his own conceptual frame and if he is to experience models of teaching appropriate to the involvement of students in thinking, the traditional lecture approach to college teaching would probably have to be largely abandoned in favor of discussion techniques (25).

The research reported here stemmed from an effort to depart from the usual pattern of social studies methods courses in order to do two things: 1. concentrate on involving the students in the examination of the assumptions (such as those about the nature of democracy) underlying the curricular decisions they would be making as social studies teachers, and 2. help prospective teachers to explore the reflective competencies that might be helpful in teaching their students to analyze public issues. This approach meant a heavy reliance on discussions with little use of the lecture method.

Not too surprisingly, perhaps, the students' reactions to such a course were mixed. Some students were quite disturbed by the course because they were not being given practical preparation for their first confrontations with students during their practice teaching which was to take place within a quarter or two. Moreover, the discussions were frustrating because ready answers were not provided to be written down in neatly organized notebooks. As one student wrote on an evaluation form, *"The class is unfair because we are tested on how well we think instead of what we know."*

Other students reacted positively to the course. They remarked that it was one of their first opportunities to think in a college course instead of simply memorizing lecture notes and underlined sentences from a textbook and then regurgitating them on tests. These same students frequently commented on their previous lack of awareness of the complexity of curricular decision-making in the social studies and on an awakened concern with the traditional history-oriented, read-and-recite pattern of social studies instruction.

It appeared that the course made demands that might evoke different reactions from students with different personality traits. Students were required to shift their conceptions of the proper instructor role. Moreover, the instructor did not provide a clear model of authority as the lecturer so often does. The analysis of assumptions underlying curricular decisions seemed to call for an openness in belief systems, and the concepts of reflective thinking that were taught emphasized a tentative approach to beliefs and the willingness to tolerate conflicting values. It seemed that the investigation of the relationship between measures of open-closed mindedness and the students' reactions to the course might be fruitful.

The Dogmatism Scale, developed by Rokeach, was *"designed to measure individual differences in the extent to which belief systems are open or closed"* (34:19). Rokeach's scale was based on earlier work in the area of authoritarianism which had produced the F-Scale (2), often called the Authoritarianism Scale. The variables measured by both scales seemed likely to be related to the reactions of students to a social studies curriculum and methods course focused on reflection. Rokeach also used the Gough–Sanford Rigidity Scale as a measure of the tenacity of individual beliefs, as opposed to the tenacity of belief systems which the Dogmatism Scale was to assess. Use of such a cognitive rigidity scale promised a more complete view of the relationships between tendencies toward lack of flexibility and openness and reactions to an inquiry-oriented course, although the Dogmatism and F-Scales were of central interest in the study.

Related Literature

There were two aspects of the student's reactions to a social studies methods course emphasizing inquiry which were of interest: 1. the relationship between dogmatism and authoritarianism and ratings of the instructor and 2. the relationship of dogmatism and authoritarianism to learning. No studies reported in the literature dealt directly with these relationships in the context of a social studies methods course. However, some related studies were of interest.

Authoritarianism and Evaluation of the Instructor. Past research makes it difficult to predict how authoritarians might react to an instructor who not only departed from the traditional content of the social studies methods course, but who also taught in an open-ended way. Obviously, the openendedness of the course might evoke negative reactions in students low in tolerance of ambiguity. Yet the findings in research relating F-Scale scores to tolerance of ambiguity have been equivocal, despite a theoretical expectation that authoritarians would be more demanding of certainty. Part of the difficulty has been the differing experimental definitions of ambiguity, ranging from reactions to lights in a darkened room (26) to the number of blanks not completed in cloze passages (17).

Studies in leadership are also of little help because of their inconclusive nature. There is some evidence that high F scorers do prefer autocratic leadership in the Navy (19), as well as leader-centered discussions in university classes (15). There are also findings which indicate that high authoritarians tend to be less critical than low authoritarians in their ratings of military leaders (19), stimulus persons in experiments (20), and parents (32). However, Maney (24) found that students' F-Scale scores were not significantly correlated with their ratings of eight instructors in a southern women's college.

Authoritarianism and Learning. F-Scale scores have generally been found to correlate negatively with intelligence scores (37). It would not be surprising, then, to find authoritarianism to be inversely related to learning. Negative correlations have been found between F-Scale scores and college grades (9, 39). Linton (23) found scholastic achievement on the provincial exams in Alberta to be inversely related to F-Scale scores. In another study (29), senior medical students who were high F scorers had more difficulty than low F scorers in learning material dealing with humanitarian philosophy and material judged to be ambiguous.

On the other hand, Bohlke (6) reported that authoritarianism was not related to learning about India or to changes in attitudes about India on the part of 225 college freshmen. And, F-Scale scores were not related to the learning of interviewing skills (3) or to the recall of nonsense sentences (33). Despite these discrepant findings, however, reviews of authoritarianism and learning (10, 16, 18) have concluded that F-Scale scores generally bear a negative relationship to conceptual learning.

Dogmatism and Evaluation of the Instructor. Research related to students' dogmatism and their evaluations of instructors is scarce. High dogmatics may find synthesis more difficult, especially under stress (21). Girault (13) reported a relatively high ($r = .45$) correlation between D scores and scores on the Walks Ambiguity Scale (intolerance of ambiguity). These factors would indicate that high dogmatic students might react negatively to a methods course that (a) required an analysis of social studies goals and a synthesis into curricular proposals on the part of the students and (b) was more ambiguous in that examinations involved the discussion of curricular issues in essay questions, rather than the testing of specific content through objective-type questions. Although studies dealing directly with dogmatism and students' ratings of instructors are not available, some findings indicate that the dogmatism of principals is related to their tendency to rate teachers as effective (28) or congenial (5).

Dogmatism and Learning. In an earlier review, Jackson and Strattner (18) indicated that dogmatism generally correlates negatively with learning. Among the learning tasks with which dogmatism has been found to be negatively related are concept formation involving a reversal of the initial task (30), paired-associate learning (1), and conditioning of verbal behavior (27). Kemp (22) found that when eighty university freshmen were given special training in problem solving and tested for gains on the Watson–Glaser Critical Thinking Appraisal, the improvement of high D scorers was not significant, while the low D scorers showed improvement significantly at the .01 level.

Frumkin (12) found that high dogmatics obtained significantly lower ($p<.01$) grades in a sociology course. Ehrlich (11), in a study that provoked a series of further research studies, found dogmatism to be significantly and inversely related to learning in an introductory sociology course, as measured by scores on a true–false test. Christensen (7) "replicated" Ehrlich's study at another university, but using students in an introductory psychology course. Using multiple-choice and essay questions as measures of learning, he did not find the inverse relationship which Ehrlich reported. He suggested that his differing results might be due to a more rigorous approach in psychology which might be more congruent with the needs of the dogmatic person. Costin (8) attempted to duplicate Ehrlich's method while using Christensen's subject matter to determine if method was a factor in the different findings. He found a slight, but insignificant relationship ($r = -.19$) between dogmatism and learning scores, and proposed that dogmatism-learning relationships are confounded by different types of dogmatism and the type of material being learned.

In order to determine if test format was a factor in the contradictory Ehrlich–Christensen findings, White and Alter (40) used both true–false and multiple-choice items with 2,099 introductory psychology students at the University of Utah. They found that with fourteen classes there were twelve negative and two positive correlations between D scores and learning scores, with a small overall correlation of $-.18$ ($p<.01$). They did not obtain different results with true–false and multiple-choice items. A study conducted at the University of Arizona with 517 introductory psychology students (41) also produced a low ($r = -.20$), but significant ($p<.001$), negative correlation between D and learning scores.

The findings related to the authoritarianism and dogmatism of students' and their ratings of instructors are scarce and inconclusive. There is some basis for expecting small, but statistically significant, negative correlations between dogmatism and authoritarianism and learning.

The Subjects and the Setting

Three sections (Group 1, N = 26; Group 2, N = 33; Group N, N = 30) of the social studies methods course were involved in the research, each taking the course a different quarter. Few women major in secondary school social studies at Utah State University. In this investigation, one group had only two females, another had four, and the third had nine. Generally, the students were juniors who had not yet done their student teaching, although an occasional senior was enrolled.

The course was a 3-hour class, meeting three times a week for a 50-minute period. There was no assigned textbook, although journal articles and sections from books were assigned for reading (36). Class meetings were typically devoted to the discussion of issues in social studies education. Lectures were presented by the instructor at no more than one out of four or five class meetings.

Because of the limited resources available, it was not possible to develop an instrument for scoring classroom behavior or to hire and train observers to score classroom behavior, although it was recognized that such information could be valuable in interpreting the findings (35). If a description of teaching style presently in the literature could be applied to the class discussions, however, it would undoubtedly be that of a socratic teacher, as compared to a recitation one (31, 35). Students were prodded to take stands on issues in social studies education and then forced to defend their stands; special attention was paid to possible inconsistencies in their beliefs and values.

In evaluating the instructor, the students were asked to indicate, *"Is there something you believe the instructor has done especially well in his teaching of the course?"* Almost without exception they responded with some variation on the theme, *"He has stimulated me to think more critically about educational issues."* Students' negative comments about the course are also instructive. When asked for "specific things you believe might be done to improve the instructor's teaching of this course," they commonly requested more specificity in assignments, lectures, and tests. This lack of specificity was a part of the attempt to set the stage for an emphasis on thinking and analysis.

Table 1.—Items from the University of Washington Survey of Student Opinion of Teaching Used in the Study

1. Interprets abstract ideas and theories clearly
2. Gets me interested in his subject
3. Has increased my skills in thinking
4. Has helped broaden my interests
5. Stresses important material
6. Makes good use of examples and illustrations
7. Has motivated me to do my best work
8. Inspires class confidence in his knowledge of subject
9. Has given me new viewpoints or appreciations
10. Is clear and understandable in his explanations

C. Is willing to give individual attention
F. Has a good sense of humor
G. Is fair in grading
H. Has presented many thought-provoking ideas
I. Has given me new tools for attacking problems
J. Shows respect for questions and opinions of students

The Data

Form 40–45 of the F-Scale (2) was used with one modification. Item No. 22, *"It is best to use some prewar authorities in Germany to keep order and prevent chaos,"* seemed historically out of context for our college samples, and was dropped from the scale. Therefore, the F-Scale means reported in this study are based on a 28- rather than a 29-item scale.

The testing instrument, which included the 28-item F-Scale, the D-Scale (Form E), and the R-Scale (34), was administered during regular class periods toward the beginning of each quarter. Administration took about 25 minutes. The reliability coefficients for the tests (split-half, corrected with the Spearman–Brown formula) for the three groups combined were: F-Scale, $r = .74$; D-Scale, $r = .82$; R-Scale, $r = .81$.

In Groups 2 and 3, the Otis Quick-Scoring Mental Ability Tests: New Edition, Gamma Test was administered as an estimate of mental ability that would consume little class time.

During one of the last two class meetings, the University of Washington Survey of Student Opinion of Teaching was administered to obtain student evaluations of the instructor. The first part of the Survey has ten items (see Table 1) on which each student rates the instructor. The instructor is to be compared with the other teachers the student has had, using the following scale: 1 = outstanding, 2 = superior, 3 = competent, 4 = fair, 5 = of less value. The numerical order of the scale was reversed in scoring the evaluations so that the correlation coefficients would be easier to interpret. Each item was included in the analysis along with a total score which was the sum of the items.

The second part of the Survey contains ten more items. The group of ten Survey items mentioned in the previous paragraph was the outgrowth of factor analyses carried out at the University of Washington (14). These items are kept constant for the University's own use of the Survey in its instructor evaluation program. The items in the second group, however, are varied occasionally in the hopes of turning up better discriminators. From the second group of items, six of interest were selected for inclusion in the analysis. These items are also presented in Table 1. Again, the individual items were summed for a total score.

Other items taken from the Survey for analysis were an estimate by the student of the grade he would receive in the course (scored A = 5, B = 4, C = 3, D = 2, E = 1), a grade for the instructor (A, B, C, D, or E) to be based on his "help" to the student in his "efforts to learn this subject," and a 3-point rating of the subject matter of the course (highly interesting, moderately interesting, not very interesting).

For each of the three groups, the students' grades in the course were used as a criterion of

Table 2.—Means and Standard Deviations on the F-, D-, and R-Scales, and the Otis Test

	F-Scale[a]		D-Scale		R-Scale		Otis	
	Mean[b]	S.D.	Mean[c]	S.D.	Mean[c]	S.D.	Mean[d]	S.D.
Group 1	88.58	17.08	149.74	23.55	85.21	14.76		
Group 2	100.03	17.74	145.70	22.89	92.16	13.65	114.45	8.82
Group 3	88.00	20.13	143.56	21.07	87.68	13.09	119.32	7.86

[a] Twenty-eight item scale (see text).
[b] With df = 2/74 and F = 3.87, the difference among the means is significant, $p<.05>.01$.
[c] Difference among the means not significant at the .05 level.
[d] With df = 2/74 and F = 3.87, the difference among the means is significant, $p<.05>.01$.

learning. In addition, for Group 2, there were available two other measures of learning: a gain score on a 3-alternative, multiple-choice test, and a gain score on the Social Issues Analysis Text No. 1 (31), aimed at determining the student's ability to identify the analytic functions indicated by statements in a discussion. In retrospect, after the completion of the course with Group 2, neither test seemed to reflect adequately the objectives of the course, so they were not administered to Group 3.

Results

Descriptive data on the three groups involved in the study are presented in Tables 2 and 3. The variability in the mean scores on the F-, D-, and R-Scales and on the Otis Test is of particular interest. The correlations between the F- and D-Scales (see Table 3) fall within the range reported by Rokeach (34:121) for his studies, although the correlations between the D- and R-Scales exceed the range of .37 to .55 reported by Rokeach (34:193). The correlations between the three personality scales and the Otis were all negative, ranging from $r = -.04$ to $r = -.23$, with none significantly larger than zero.

Means and standard deviations for the students' evaluations of the instructor on the University of Washington Survey of Student Opinion of Teaching are contained in Table 4. It will be recalled that a 5-point rating scale was used, scored so that the highest rating received a value of 5 and the lowest a value of 1. The rating of interest in the subject matter of the course was on a 3-point scale.

Generally, the mean ratings are at or above the median scale value. Table 5 provides information for testing the instructor's impression that students tended to have rather strong positive or negative reactions to the courses. Although the distributions of ratings are not bimodal, they do tend to be platykurtic, with extreme scores extending the distributions. The shapes of the distributions lend support for the conjecture that reaction to the methods course was diverse.

Table 3.—Correlations Among the F-, D-, and R-Scales

	F-Scale	D-Scale
1. F-Scale		
2. D-Scale	.70* .61 .73	
3. R-Scale	.49 .46 .58	.27 .62 .76

*Coefficients listed in descending order Group 1 first; Group 2, second; Group 3, third.

Item intercorrelations for both parts of the Survey were generally moderate and positive, indicating relationship but considerable independence among the items (37).

The correlations of individual items with the total score for the first part of the Survey were relatively high (from $r = .41$ to $r = .86$); the item-total correlations were not so consistently high for the second part of the survey, raising some question about the validity of that total score. In addition, the low correlations between the total scores for the two parts ($r = .39$, .42, and .19) indicated that the two total scores should not be pooled as one variable. A reflection of the low relationship between the two total scores was the relatively high correlation of the first part total with the grade given the instructor ($r = .43$, .65, .65) as contrasted with the lower correlations for the total of the second group of items ($r =$ .35, .25, —.04).

What of the relationships between F-, D-, and R-scores and student reactions to the course? Because there were few correlations of even a moderate magnitude, and only seven of these were significant, the matrix of coefficients for the personality measures and the Survey items is not presented here. The total score for the first ten items (which appeared to be the more reliable and valid of the two total scores) was not correlated significantly ($p<.05$) with the F-, D-, R-, or Otis-scores.

There were only three significant correlations for the F-Scale. Two of these—for Groups 1 and 3—were with Item C, *"Is willing to give individual attention"* ($r = -.61$ and .40). The third was with the total score for the second part of the Survey ($r = -.55$). Dogmatism scores correlated significantly with only one item—Item 3, *"Has increased my skills in thinking"* for one group ($r = .34$). Gough Sanford Rigidity scores were involved in two significant correlations—identical in numerical value and opposite in sign. The positive correlation was with Item 4, *"Has helped broaden my interests"* ($r = .51$), and the negative correlation was with Item F *"Has a good sense of humor"* ($r = -.51$).

There does appear to be a rough correspondence between fluctuations in the means for the rating scale totals and the grade assigned to the instructor (Table 4) and the decrease in the D-Scale means (Table 2) from Group 1 to 3. This rough association must be interpreted cautiously, of course. A similarly negative association between mean course grades (Table 6) and dogmatism means is noticeable. It may be that the

Table 4.—Means and Standard Deviations for the University of Washington Survey Scores

	Group 1		Group 2		Group 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Item - 1 (Interprets ideas)	3.42	.84	3.60	.86	3.88	1.01
Item-2 (Gets me interested)	3.16	1.01	2.91	1.07	3.60	.86
Item-3 (Skills in thinking)	3.68	.82	3.64	.99	3.84	.75
Item-4 (Broaden interests)	3.17	.71	3.24	1.12	3.48	1.19
Item-5 (Important material)	3.42	.61	2.88	.78	3.52	.96
Item-6 (Examples and illustration)	3.63	1.01	3.15	1.00	3.88	.93
Item-7 (Motivated me)	2.68	.58	2.61	1.02	2.92	1.04
Item-8 (Confidence in knowledge)	3.32	.94	3.54	1.06	3.75	.94
Item-9 (New viewpoints)	3.68	.88	3.82	.95	3.92	.90
Item-10	3.21	.79	3.00	.97	3.42	.97
Total	33.05	4.58	32.39	6.65	36.44	7.19
Item-C (Individual attention)	3.58	.96	3.58	1.00	3.96	1.17
Item-F (Sense of humor)	3.68	.67	3.48	1.00	4.04	.84
Item-G (Fair in grading)	2.68	.94	2.87	.76	2.96	1.16
Item-H (Thought-provoking)	3.63	.90	3.52	1.15	4.08	1.00
Item-I (New tools)	3.32	.82	3.42	.97	4.00	.95
Item-J (Respect for students)	3.32	1.00	3.54	1.03	3.88	.83
Total[a]	20.21	3.01	20.24	4.45	25.21	10.66
Instruc. Grade	3.94	.72	3.91	.73	4.04	1.12
Subject Matter	2.16	.60	2.16	.73	2.43	.66
Expected Grade	3.50	.51	3.63	.67	3.57	.60

[a] With df = 2/73 and f = 4.35, the difference among the means is significant, $p<.05>.01$.

trend toward higher ratings of the instructor is due not to a decrease in dogmatism, but to a more favorable classroom climate with a group of students who receive higher grades.

If one accepts variability in course grades as a criterion of learning, the coefficients reported in

Table 5.—Distribution of Ratings of Instructor [redacted] University of Washington Survey—Totals for Two Groups of Items

Total Rating	Survey Total 1			Groups Combined	Survey Total 2			Groups Combined
	G-1*	G-2	G-3		G-1	G-2	G-3	
55			1	1				
54								
53								
52								
51								
50								
49								
48								
47			1	1				
46								
45		1	1	2				
44			1	1				
43	1			1				
42		2		2				
41			3	3				
40		2	2	4				
39	1	3	2	6				
38	2	2	2	6				
37			1	1				
36	2	2	2	6				
35	1	1		2				
34	1	1	2	4				
33	2	2	1	5				
32	2	2	2	6				
31	3		3	6				
30	1	2	3	6			1	1
29		3		3		1	2	3
28	1	1		2			3	3
27	2	2		4				
26	1	2	1	4			1	1
25		1		1	1	2	3	6
24		1		1	2	5	3	10
23					2	1	2	5
22		1		1	2	7	2	11
21					1	2	1	4
20					5	2	1	8
19					1	1		2
18					1	4	1	6
17		1		1	1	1		2
16					1			1
15					2	2	1	5
14								
13							2	2
12						2		2

* G-1 stands for Groups 1, G-2 for Group 2, etc.

Table 6.—Means and Standard Deviations for Course Grades, Multiple Choice Gain Scores, and SIAT No. 1 Gain Scores

	Group 1		Group 2		Group 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Course Grades	3.11	.57	3.15	.87	3.40	.86
Multiple Choice Gain			5.00	1.98		
SIAT No. 1 Gain			5.75	2.46		

Table 7 reflect the relationships one might expect between measures of learning and Otis IQ scores (with the exception of the Multiple Choice Gain Score, which is of doubtful validity). There are, however, no significant correlations (not even any sizable ones) between the F-, D-, and R-Scale scores and the three learning criteria.

Conclusions

The scattering of significant correlations between F-, D-, and R-scores and students' ratings of the instructor without a discernible pattern hardly gives credence to the notion that there is a consistent relationship between students' open-closed mindedness and their evaluations of the instructor of an inquiry-oriented social studies methods course. The complete absence of significant correlations between the three personality variables and the measures of learning for the courses indicates no support for the idea that authoritarianism, dogmatism, or cognitive rigidity are related to learning in such a course.

Of course, one can always say "if" we had used a different instructor rating form or different measures of learning, significant relationships might have emerged. However, the measures are of the kind typically used to assess students' learning and to obtain ratings of instructors. (In fact, the University of Washington Survey is probably a much more carefully developed measure than many teachers use in assessing student opinions of their teaching.) With these "typical" instruments, few significant relationships emerged using three different groups of students at three different points in time.

It might also be worth noting that many of the correlations would have been statistically significant had the size of the samples been large—say a total of five hundred. The literature on the relationship between personality and learning has been cluttered with statistically significant but educationally trivial findings. (Examples are the White and Alder (40) and Zagona and Zurcher (41) studies.) The important question is how much variance in learning or in ratings of the instructor can be accounted for by variability on the personality measures. The use of reasonably sized, but relatively small groups saves the practitioner who is interested in practical application, not theoretical implications (as important as the latter may be in the context of scientific investigation) from having to evaluate seemingly unimportant relations. Practically, our results indicate little relationship between open-closed mindedness and reactions to an inquiry-oriented social studies method course.

Table 7.—Correlations Among Course Grades, Multiple Choice Gain, SIAT No. 1 Gain, Otis IQ, and F-, D-, and R-Scale Scores

	IQ	F-Scale	D-Scale	R-Scale
Course Grades	——*	-.01	.08	.10
	.48[b]	-.26	-.03	-.04
	.49[a]	-.03	.03	.14
Multiple Choice Gain	——	——	——	——
	-.03	-.17	-.24	-.23
	——	——	——	——
SIAT No. 1 Gain	——	——	——	——
	.40[a]	-.17	-.04	.02
	——	——	——	——

* Coefficients in descending order—Groups 1, 2, and 3.

[a] Significant at the .05 level.

[b] Significant at the .01 level.

FOOTNOTE

1. The research reported here was carried out pursuant to a contract with the U.S. Office of Education.

REFERENCES

1. Adams, H. E.; Vidulich, R. N., "Dogmatism and Belief Congruence in Paired-Associate Learning," *Psychological Reports*, 10:91–94, 1962.

2. Adorno, T. W.; Frankel-Brunswik, E.; Levinson, D. J.; Sanford, R. N., *The Authoritarian Personality*, Harper and Brothers, New York, 1950.

3. Balensky, B., "Relationship Between the Ability to Acquire Interviewing Skills and Authoritarianism and

Manifest Anxiety," *The Personnel and Guidance Journal*, 42:40–42, 1963.

4. Ballinger, S., "The Social Studies and Social Controversy," *School Review*, 71:97–111, 1963.
5. Bentzen, M. M., "Some Effects of Dogmatism in Elementary School Principals and Teachers," paper presented at the meeting of the American Educational Research Association, Chicago, February 1966.
6. Bohlke, R. H., "Authoritarianism and Attitudes of College Students Toward India," *Journal of Educational Sociology*, 34:145–159, 1960.
7. Christensen, C. M., "A Note on 'Dogmatism and Learning,'" *Journal of Abnormal and Social Psychology*, 66:73–75, 1963.
8. Costin, F., "Dogmatism and Learning: A Follow-up of Contradictory Findings," *The Journal of Educational Research*, 59:186–188, 1965.
9. Davids, A., "The Influence of Ego-Involvement on Relations Between Authoritarianism and Intolerance of Ambiguity," *Journal of Consulting Psychology*, 20: 179–184, 1956.
10. DiVesta, F. J., "Meaningful Learning: Motivational, Personality, Interpersonal, and Social Variables," *Review of Educational Research*, 31:511–521, 1961.
11. Ehrlich, H. J., "Dogmatism and Learning," *Journal of Abnormal and Social Psychology*, 62:148–149, 1961.
12. Frumkin, R. M., "Dogmatism, Social Class, Values, and Academic Achievement in Sociology," *Journal of Educational Sociology*, 34:393–403, 1961.
13. Girault, E. S., "The Relationship Between Teachers' Dogmatism and Teachers' Concept of the Discipline of History," paper presented at the meeting of the American Educational Research Association, Chicago, February 1968.
14. Guthrie, E. R., *The Evaluation of Teaching: A Progress Report*, University of Washington, Seattle, 1954.
15. Harmon, F. S., "A Measurement of Authoritarian Attitude Toward Discussion Leadership," *Quarterly Journal of Speech*, 41:140–144, 1955.
16. Harvey, O. J., "Authoritarianism and Conceptual Functioning in Varied Conditions," *Journal of Personality*, 31:452–470, 1963.
17. Honigfeld, G.; Platz, A.; Gillis, R. D., "Verbal Style and Personality: Authoritarianism," *Journal of Communication*, 14:215–218, 1964.
18. Jackson, P. W.; Strattner, N., "Meaningful Learning and Retention: Non-Cognitive Variables," *Review of Educational Research*, 34:513–529, 1964.
19. Jones, E. E., "Authoritarianism as a Determinant of First-impression Formation," *Journal of Personality*, 23:107–127, 1954.
20. Kates, S. L., "First Impression Formation and Authoritarianism," *Human Relations*, 12:277–286, 1959.
21. Kemp, C. G., "Critical Thinking: Open and Closed Minds," *The American Behavioral Scientist*, 5:10–14, 1962.
22. Kemp, C. G., "Improvement of Critical Thinking in Relation to Open-Closed Belief Systems," *Journal of Experimental Education*, 31:321–323, 1963.
23. Linton, T. E., "Achievement and Authoritarianism as Components of Socio-Cultural Factors," *Alberta Journal of Educational Research*, 13:129–134, 1967.
24. Maney, A. C., "The Authoritarianism Dimension in Student Evaluations of Faculty," *Journal of Educational Sociology*, 32:226–231, 1959.
25. McKeachie, W. J., "Research on Teaching at the College and University Level," in Gage, N. L. (ed.), *Handbook of Research on Teaching*, Rand McNally, Chicago, 1963, pp. 1126–1127.
26. Millon, T., "Authoritarianism, Intolerance of Ambiguity, and Rigidity under Ego and Task-involving Conditions," *Journal of Abnormal and Social Psychology*, 55:29–33, 1957.
27. Moore, D. M., "Dogmatism and Anxiety in the Conditioning of Verbal Behavior," *Dissertation Abstracts*, 22:2468, 1962.
28. Musella, D., "Open- and Closed-mindedness as Related to the Rating of Teachers by Elementary School Principals," *Journal of Experimental Education*, 35:75–79, 1967.
29. Neel, A. F., "The Relationship of Authoritarian Personality to Learning: F-Scale Scores Compared to Classroom Performance," *Journal of Educational Psychology*, 50:195–199, 1959.
30. Ohnmacht, F. W., "Effects of Field Independence and Dogmatism on Reversal and Nonreversal Shifts in Concept Formation," *Perceptual and Motor Skills*, 22: (no. 2) 491–497, 1966.
31. Oliver, D. W.; Shaver, J. P., *Teaching Public Issues in the High School*, Houghton Mifflin, Boston, 1966.
32. Phares, E. J., "Note on Authoritarianism and Attitudes Toward Parents," *Journal of Clinical Psychology*, 16:389–390, 1960.
33. Pyron, B.; Kafer, J., Jr., "Recall of Nonsense and Attitudinal Rigidity," *Journal of Personality and Social Psychology*, 5:463–466, 1967.
34. Rokeach, M., *The Open and Closed Mind*, Basic Books, New York, 1960.
35. Shaver, J. P., "The Ability of Teachers to Conform to Two Styles of Teaching," *Journal of Experimental Education*, 32:259–267, 1964.
36. Shaver, J. P.; Berlak, H., *Democracy, Pluralism, and the Social Studies: Readings and Commentary*, Houghton Mifflin, Boston, 1968.
37. Shaver, J. P.; Richards, H. E., *Open-Closed Mindedness of College Students in Teacher Education*, ERIC # ED-026303, report of U.S.O.E. Project No. 7–8056, Utah State University, Logan, 1968.
38. Smith, P. G., "How We Think: A Re-examination," *The Educational Forum*, 31:411–420, 1967.
39. Stott, S. S., "A Comparison of the Authoritarian Personality F-Scale With The Allport–Vernon–Lindzey Study of Values," unpublished master's thesis, University of Utah, Salt Lake City, 1954.
40. White, J.; Alter, R. D., "Dogmatism and Examination Performance," *Journal of Educational Psychology*, 58: 285–289, 1967.
41. Zagona, S. V.; Zurcher, L. A., Jr., "The Relationship of Verbal Ability and Other Cognitive Variables to the Open-Closed Cognitive Dimension," *Journal of Psychology*, 60:213–219, 1965.

● In elementary grades, every teacher is a teacher of reading, but the typical student receives no further training in reading after he leaves elementary school. Yet all students can benefit from instruction in advanced reading skills. Most high schools have just a small corps of reading specialists who work only with retarded readers. This is inadequate. Offering reading instruction only to the retarded reader is like offering physical education only to the physically handicapped.—Lee H. Mountain in *NJEA Review.*

FIELD NEWS

Willard E. North
Central Missouri State College

THE REGIONAL EDUCATIONAL LABORATORY for the Carolinas and Virginia (RELCV) is now the National Laboratory for Higher Education (NLHE). The laboratory was established in 1965 and has concentrated on higher education, devoting a majority of its budget to programs to improve instruction and administration at 2-year and 4-year colleges. It is the only federally funded education laboratory focusing on these areas in higher education.

The laboratory's elementary and secondary school unit has been renamed the Center for Individualized Instructional Systems (CIIS) and will continue to function as a division of the laboratory.

The NLHE Junior and Community College Division has conducted teacher training workshops for faculty members at eighty-nine colleges and universities in twenty-three states, Canada, and Puerto Rico. About fifteen hundred faculty members are now individualizing their instruction using techniques learned at the workshops.

Computer programs to improve administrative efficiency, developed by the laboratory's senior college division, are now in use at seventy colleges and universities in thirty-three states and Canada. Also, the division has provided special training in management techniques and in instructional improvement to faculty members and administrators at more than a score of 4-year institutions. The laboratory's innovative elementary school curriculum, the Individualized Mathematics System (IMS), will be used by about 75,000 pupils in thirteen states and the District of Columbia during the 1971–1972 school year.

According to Everett H. Hopkins, NLHE President, the name change does not reflect a dramatic shift in the laboratory's organization or operations. Although the laboratory's budget has steadily increased to match expanding operations over the past few years, its main mission is to develop "exportable educational products" that can be used at any school or college with a minimum of orientation and training. The laboratory will continue to develop products to meet the needs of the schools and colleges in the Carolinas and Virginia. It is the intent of the laboratory to make available successful products to any institution regardless of its location.

* * * * * *

AS WITH MOST institutions of higher education, the most widely used teaching methodology at Michigan State University is one involving lectures, text, and recitation sections. This time-honored method is often used because of economic or administrative constraints rather than learning or pedagogical considerations. Given the practical necessity for using the lecture-text teaching model, M. Ray Denny and Stanley C. Ratner have conducted a study exploring ways which improve student learning. This project was supported by a grant from the Educational Development Program at Michigan State.

Subjects used in the study were enrolled in Psychology 200, a large enrollment course (350 students per term) noted for its high information content and demanding nature. The student motivation and morale sometimes became a problem due to the large lecture format. In an attempt to remediate these problems and improve student learning, Drs. Ratner and Denny introduced three innovations to the course: (1) offering optional, non-graded, self-scored "Friday quizzes" covering the week's lecture and text assignments; (2) permitting students receiving a midterm grade of 2.0 or below to take the exam again (with a maximum grade of 2.0 awarded); (3) tape recording all lectures and making the tapes available to students in the main Michigan State University library.

Optional quizzes were designed to give students practice taking tests, to provide feedback on the adequacy of the previous week's learning, and to help student's identify important course concepts. By offering optional quizzes throughout the term, students were able to estimate their mastery of specified blocks of subject matter and therefore, assume greater responsibility for their own learning. Quizzes were given on Friday so students opting not to take the quiz had an early weekend. Each quiz was 15 minutes in length, typically requiring students to fill in key words in paragraphs describing important concepts. Grading was done by students after the exam, so feedback was immediate. Wrong answers were discussed immediately under the auspices of qualified graduate teaching assistants, thus providing additional instruction in the most difficult concepts. Students who did not score at least 80 percent were encouraged to take a second "equivalent" exam over the same topics. If they achieved less than 80 percent on the second try, they could take a third quiz. In no case was it obligatory to take any quiz, and whatever the score, it did not count toward the student's final grade. To evaluate these innovations, records were kept on tape utilization, midterm retakes, percent of students participating in quizzes, and quiz scores. Informal attitude measures were taken and correlations run to determine the relationship between quiz participation, quiz scores, and final course grades. Student utilization of tapes and quizzes steadily increased throughout the year often achieving 90 percent participation in the optional quizzes. Approximately 30 percent of the students chose to retake the midterm exams as a source of additional practice and a way to improve learning. The most rewarding result, however, was the almost linear positive relationship between achievement and optional quizzes (non-graded) and final course grades (based on midterm and final exam scores). In other words, the more practice students had on taking optional tests, the clearer became the course objectives, more feedback was received on the adequacy of their learning, and study activities could be adjusted with the resultant increase in learning as demonstrated by greater achievement on exams. The authors feel that several unplanned benefits also have occurred from the use of optional quizzes, for example, the student/faculty interpersonal relationships have improved by virtue of the informal discussion following each 15 minute quiz, and a sense of team work leading to mastery of course content was evolved.

* * * * * *

THE GROWTH OF SIMULATION GAMES, the type and coverage of available games, cost and use of games in education and training, and a brief history of simulation games are all covered in "What Is It You Want To Know?" an article by David Zuckerman and Robert E. Horn. Zuckerman and Horn are co-authors of the "Guide to Simulation Games for Education and Training," a recently published catalog of four hundred games and simulations. The free article is available from Information Resources, Incorporated, 1675 Massachusetts Avenue, Cambridge, MA 02138.

* * * * * *

THE ADULT EDUCATION Research Conference for 1972 will be held in the LaSalle Hotel, Chicago, Illinois, April 7–9. Those interested are invited to submit papers, symposium, and workshop proposals.

Papers are typically empirical, theoretical, or methodological reports with implications for adult education. Focus is on design and research method rather than on findings. A one page abstract focusing on objectives, design, and methods of study must be submitted by November 15, 1971.

Symposia are devoted to explication of theoretical, methodological, or other issues. Tentative proposals listing topics, possible resource people, organizers, and potential participants are due by November 15.

Workshops are to be a new feature of the 1972 conference. They are envisioned as day-long experiences centered on a current methodological issue or problem of practice. A featured speaker or panel will delineate the problem or issue. Workshop participants will then engage in further study, discussion, or work sessions. Proposals, listing the suggested topic, resource persons, participants, and general design of the workshop, are due by November 15.

Abstracts, proposals, and requests for information should be sent to: Prof. Robert Bruce, Department of Education, Cornell University, 112 Stone Hall, Ithaca, NY 14850.

* * * * * *

A RESEARCH TEAM AT PURDUE University has developed a new series of tests that they hope will help in the assessment of children's intellectual development regardless of social-economic, cultural, and racial differences. Use of the new tests with Gary, Indiana, school children indicates that the measurements yield valid results for pupils of various backgrounds. Children from both low and middle income areas were tested. The group included blacks, whites, and children from Spanish speaking homes. More than 14,000 children in the second, fourth, and sixth grades took part in the testing.

The team sought to produce tests which would measure ability but would be less dependent on school-related skills than the tests typically now in use. To minimize the need for such skills, pupils receive test directions through the use of movies, slides, and audio tapes. Test booklets make extensive use of cartoons, geometric shapes, and other visual forms.

Prof. Robert B. Kane, one of the investigators, believes the use of films and tapes contributes not only to the "cultural fairness" of the test but may also motivate some pupils to try harder because of the novelty of getting away from paper and pencil type tests.

The test attempts to measure five factors of ability: concept formation, language development, logical thinking, problem solving, and response style. It can be administered to groups of children and yields thirty-five scores or sub-scores. It is anticipated that the main use of the test will be to supplement regular intelligence and achievement measures given in the schools.

The project began in mid-1970 and was supported by a grant of $111,000 from the U. S. Office of Education. The staff of Purdue's Educational Research Center and specialists in child development, education, and psychology worked in the program. Other key personnel on the project are J. William Asher and John F. Feldhusen, Professors of Education and Psychology; Gerald E. Gruen, Associate Professor of Psychology; Ernest McDaniel, Professor of Education; M. Irene Stevens, Instructor of Child Development; John O. Tyler, Associate Professor of Education; and Grayson H. Wheatley, Assistant Professor of Education and Mathematics.

* * * * * *

DR. BERNARD SPOLSKY, Linguistic Researcher at the University of New Mexico, is carrying out a study to find a more effective way to teach English to 6-year-old Navajo Indians. The investigator believes that young Navajos (90 percent of whom reach first grade knowing no English) would learn English better if they were taught to read and write their native language first.

The linguist trained adult Indians to conduct tape-recorded interviews with more than 150 children of the 120,000 member Navajo nation to determine what instructional approach should be taken in the children's native tongue. Each interview has been punched into IBM cards and subjected to analysis, using the University of New Mexico's IBM system/360, Model 67. The analysis produced an alphabetical index of all the words used in the interviews, noting for each the sentence in which it was used and the frequency of use. The index is 2,673 pages long.

Future analysis will produce lists of Navajo words in various letter combinations—providing clues to the children's perception of letter tandems, such as vowels and diphthongs. Investigators also intend to use the computer to determine the children's knowledge of grammatical parts such as verb forms and word roots.

The investigators plan to develop texts that could be substituted for the Dick-Jane-Spot type of text now currently used by most Navajo children. Their first instruction could then be in Navajo and could feature Indian people in the stories. The investigator feels that Navajo children could relate much more readily to this type of material than to that appearing in the traditional texts.

* * * * * *

A NEW TECHNIQUE DESIGNED to promote agreement by key groups on college goals is being developed and tested by the National Laboratory for Higher Education (NLHE). A principle feature of the technique is the use of a questionnaire method to avoid face to face confrontations. This procedure usually involves administering the questionnaire to representatives from an institution's administration, faculty, student body, board of trustees, alumni, civic leaders, and other key groups. These representatives rank college goals by priorities, and then get a summary showing priorities set by each group. The questionnaire is administered three times, and participants who differ from group ratings cite reasons for their disagreement.

Preliminary testing of the procedure was carried out during 1970 at five southern colleges and universities. In these trial runs, key groups at each institution often move from broad disagreement on the first questionnaire to nearer accord by the time the final questionnaire is completed. NLHE now is developing techniques that will aid college and university planning groups in deriving short-range objectives from the goals and in adopting "management by objectives" methods to reach objectives.

Inquiries about the procedure should be sent to the National Laboratory for Higher Education, Mutual Plaza, Durham, NC 27701.

Readers are invited to send items of research interest to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri 64093

Among Those Present

NORMAN J. BAUER is Professor of Education and Chairman of the Department of Curriculum and Instruction at the College of Arts and Science, State University of New York, Geneseo. He received his BS at Wisconsin State University in Oshkosh, his MA from Northwestern University, Evanston, Illinois, and his EdD from Indiana University. His background includes teaching in junior and senior high schools and at Eastern Illinois University, Charleston, and Directing the Laboratory School at Wisconsin State University, Oshkosh. His articles and reviews appear in a number of professional journals.

EDWIN L. HERR is Professor of Education and Chairman, Graduate Programs in Counselor Education, The Pennsylvania State University. His most recent prior employment was as Director, Bureau of Guidance Services, Pennsylvania Department of Public Instruction and Associate Professor of Education, State University of New York at Buffalo. He received the BS from Shippensburg State College and the MA and EdD from Teachers College, Columbia University.

RICHARD A. MEADE is a Professor of English Education in the School of Education of the University of Virginia at Charlottesville and is Chairman of the Department of Curriculum and Instruction there. He has just completed a 6-year term on the Executive Committee of the Conference on English Education and in 1967 ended a term of similar length on the Curriculum Commission of the National Council of Teachers of English. Mr. Meade received his PhD in Education and English at the University of Virginia in 1941. W. GEIGER ELLIS is Assistant Professor of English Education in the College of Education of the University of Georgia. He has taught and supervised in public schools for 10 years. For the past few years he has been an evaluator of documents for the Educational Resources Information Center. He has published articles in various educational journals and has taught in or directed five federally funded institutes, the last of which was held in the Philippine Islands. In 1967 he received his EdD in English Education at the University of Virginia.

AGNES G. REZLER is an Associate Professor at the Center for Educational Development, University of Illinois College of Medicine. From 1962 to 1967 she taught educational psychology at De Paul University in Chicago and at Northern Illinois University in De Kalb. Her degrees were received from Hunter College, New York City, BA; Stanford University, California, MA; and University of Chicago, PhD. ALEXANDER S. ANDERSON holds an MD from the University of Michigan Medical College and an MA in education from the University of Illinois in Urbana. From 1966 to 1969 he was Chief, Training Section, at the Center for Educational Development, University of Illinois College of Medicine. Presently he is Director of the Continuing Health Education Council of the Regional Medical Program in Hawaii.

BERNARD J. SHAPIRO, Assistant Professor of Education at Boston University, received the EdD from Harvard University and is specifically concerned with problems in curriculum evaluation and cognitive development. ROBERT E. WILLFORD, Director of Educational Programs at the Radiant Corporation, received the EdD from Arizona State University. PHYLLIS P. SHAPIRO, Assistant Professor of Education at Emmanuel College, received the EdD from Boston University. Her area of special interest is the language arts.

JAMES P. SHAVER is Professor of Educational Research and Chairman of the Bureau of Educational Research at Utah State University. He received a BA from the University of Washington and AMT and EdD degrees in social studies education from Harvard University. His doctorate was in the research in instruction program. HYRUM E. RICHARDS received his PhD in educational psychology from Utah State University after obtaining a BS from Brigham Young University and a master's in psychology. He is presently Director of Special Services, Casper–Midwest Schools, Casper, Wyoming.

FLORENCE SHEPPS received her bachelor's degree from Long Island University. After several years of elementary and junior high school teaching, she undertook intensive graduate study at Queens College, City University of New York, completing her master's degree in 1969. Her husband, a psychologist interested in school attitude problems, cooperated in the re-analysis of her research data in order to clarify the differences in school attitude between boys and girls. RONALD SHEPPS, who holds graduate degrees from Yale and Case Western Reserve Universities, maintains a private practice in Kew Gardens Hills, Queens.

DAVID P. YENS is a graduate assistant at the Computer Assisted Instruction Laboratory of the Pennsylvania State University where he is completing his dissertation for a PhD in Educational Psychology. He earned his AB from Boston University and an MEd from the University of Pittsburgh. Prior to attending Penn. State, Mr. Yens was an officer in the U. S. Army Corps of Engineers and was with the American Institutes for Research in Pittsburgh where he worked with programmed instruction and computer-assisted instruction systems design.

GUIDELINES FOR J. E. R. CONTRIBUTORS

The Journal of Educational Research publishes research articles and critiques designed to advance the scientific study of education and improve field practice. Appropriate articles are usually relatively short, within a range of 1,500 to 3,000 words.

REQUIRED ABSTRACT

Each *JER* manuscript is to be accompanied by a short abstract which capsules the nub of the article. It will be published at the beginning of the article and in many cases can take the place of a summary at the end.

Your abstract should conform to the following standards:

1. In a research paper, include statements of (a) the problem, (b) the method, (c) the data, and (d) the conclusions.

2. In a review or discussion article, state the topics covered and the central thesis (es).

3. Do not exceed 120 words. You may use standard abbreviations freely, provided the meaning is clear. Use complete sentences. Do not repeat information in the title.

TEXT

Articles generally should possess certain elements of structure in a clearly organized presentation. The ordering of these elements is not intended to be prescriptive:

Statement of the Problem. State clearly the nature, scope, and significance of the problem.

Related Research. Presentation and discussion of related research should be minimal and highly selective. In general, you should include only those references essential to clarify the problem under examination.

Population and Methodology. Adequate description of population and methodology is essential.

Presentation of Data. When concise, clearly organized and properly captioned tables, figures, and charts are used to supplement verbal presentation of data, titles of tables should be self-explanatory, and discussion of graphic data should be pointed primarily toward interpretation, omitting mere recapitulation.

Means and standard deviations should be reported. Assumptions of statistical tests should be discussed. Data on reliability and validity of the measures should be included.

Findings and Conclusions. The analysis of the data should extend beyond mere presentation to state precisely what has been determined as a result of the study. Rival hypotheses (or threats to the validity of the study) should be presented when applicable.

Implications for Practice. You are encouraged to go beyond your data to suggest implications for field practice. These may be alternative suggestions related to alternative conclusions of your research.

Summarizing Statements. A summary statement of conclusions and/or interpretive discussion may supplement the abstract.

STYLE

Certain suggestions are offered here to achieve uniformity in publication style and to conserve the time and energy of the author and editorial staff in the preparation of copy. Obviously, your article must be grammatically correct and consistent in style. Strive for simplicity and clarity of expression. *The Journal* follows *A Manual of Style,* 12th ed., Chicago, University of Chicago Press, 1960.

Typed Copy. Copy should be double-spaced with wide margins. Submit only the original typed copy. The editors will accept dittoed or mimeographed copy if it is legible.

List of References. Only references cited in the article should be included—in an alphabetical, numbered list of *References* at the end of the article. References should be complete as to author's name, title of book or article, place of publication, publisher or name of magazine, date of publication, and page numbers. ***Do not abbreviate titles of magazines.***

References Cited in Text. Refer to references by arabic numerals in parentheses, as *(3)*. If a specific page reference is intended, use the reference number, colon and page number as *(3:116–7)*. Please examine recent issues of the JER for style and form.

Footnotes. In general, avoid explanatory footnotes and incorporate the content of the footnote in the text. If footnotes are absolutely necessary, list them in sequence in a section entitled *Footnotes* preceding the list of *References,* and designate each in the text by a superscript number as *book.*[2]

Subheads. Articles are frequently improved by the judicious use of subheads. However, avoid the use of the title, *INTRODUCTION,* for a lead section.

Tables. Prepare any tables precisely as they are to appear in *The Journal,* caption each with a title, and number consecutively with arabic numerals: *Table 3. STUDENT COSTS AT MIDLAND UNIVERSITY.*

Figures. Draw any graphs and charts in black ink on good quality paper, title each, and number consecutively thus: *Figure 4. SCHOOL ENROLLMENT.* Data should not be framed, and ordinates and abscissas should be appropriately shortened to occupy as little space as possible.

Tables and figures. Tables and figures must be original copies acceptable for reproduction. A charge will be assessed for any redrawing or re-typing of tables or figures.

***Technical Symbols.* All symbols, equations, and formulae must be very clearly represented, especially when written in by hand.**

COSTS

The publisher charges a contributor's fee of $6 per printed page of approximately 1,200 words, billed upon publication. Authors are charged for changes in tables, figures, or copy made when article is in galley form. Each contributor receives 10 complimentary copies of the issue in which his article appears. Reprints are charged at cost. A price schedule will be sent to each contributor.

PROOFREADING

We will send you proofs for correction (with instructions for handling). Any major changes made in the proofs that were not incorporated in your original copy will be an added expense to you. (Errors that we make, naturally, will be at our expense.)

Before submitting your MS, please check your article according to the items outlined on this page. Keep an exact carbon copy of the manuscript you submit so that in case a question arises the editors can refer you to specific pages, paragraphs, or lines for clarification by letter.

OPTIONAL "RESEARCH BRIEFS"

In lieu of submitting a full-length manuscript, you may wish to submit an extended note for publication in the *Research Briefs* section of *The JER.* Such a note should conform to the standards for "abstracts" outlined earlier in these Guidelines, except that it may be as long as 500 words. There is no contributor's fee associated with the publication of a *Research Brief.*

Dembar Educational Research Services, Incorporated, Box 1605, Madison, Wisconsin 53701